**STAT 101L: Data Analysis and Statistical Inference**

Fall 2021

MW 12-1:15 PM in Social Sciences 139

*Please also check when your lab session will take place.*

**Instructor:** Jacob Smith (jacob.f.smith@duke.edu)

**Office Hours in Person:** Monday and Wednesday 1:30 PM to 2:20 PM and by appointment. These will be held outside of the Old Chem Building, I will either be on the porch right at Old Chem or at one of the tables outside the Perkins Café.

**Office Hours on Zoom:** Tuesday from 5:00 to 6:20 PM, please sign up on Calendly.

* Zoom Link: <https://duke.zoom.us/j/95400115230>.
* Sign Up Page: <https://calendly.com/jacobfhsmith>. You may sign up for up to two consecutive meeting slots. I will invite you into the meeting from the waiting room if you have signed up.

I welcome you to come to office hours to discuss assignments, but also if you would like to discuss your experience as a college student at Duke, something in the news, or anything else.

**Head TA:** Joseph Mathews ([joseph.mathews@duke.edu](mailto:joseph.mathews@duke.edu)).

**Office Hours:** Wednesdays 4-6 PM (In Person)

**Office Hours for Other TAs- Passcode is Welcome!**

|  |  |  |
| --- | --- | --- |
| TA | Office Hours | Zoom Room Link or Location |
| Joseph Ekpenyong ([joseph.ekpenyong@duke.edu](mailto:joseph.ekpenyong@duke.edu)) | Thursday 9-11 AM | In Person: TBD |
| Simone Nabors ([simone.nabors@duke.edu](mailto:simone.nabors@duke.edu)) | Wednesday 2:30-4:30 PM | Zoom: <https://duke.zoom.us/j/95108574551>? |
| Ben Wallace ([ben.wallace@duke.edu](mailto:ben.wallace@duke.edu)) | Monday 7-9 PM | Zoom: |
| Ginny Hagerty ([virginia.hagerty@duke.edu](mailto:virginia.hagerty@duke.edu)) | Tuesday 9:30-11:30 AM | In Person: TBD |
| Aeva Jazic ([aeva.jazic@duke.edu](mailto:aeva.jazic@duke.edu)) | Monday 5-7 PM | Zoom: <https://duke.zoom.us/j/7509412275> |
| Gaurav Sirdeshmukh ([gaurav.sirdeshmukh@duke.edu](mailto:gaurav.sirdeshmukh@duke.edu)) | Friday 2-4 PM | In Person: |

*You are welcome to go to the office hours of any TA.*

**Lab Sections**

|  |  |  |
| --- | --- | --- |
| **Section** | **Graduate TA** | **Undergraduate TA** |
| Thurs 12:15 PM (Link 087) | Joe Mathews | Ben Wallace |
| Thurs 3:30 PM (Link 087) | Joe Mathews | Simone Nabors |
| Thurs 5:15 PM (Link 087) | Joe Ekepenyong | Ginny Hagerty |

*All labs will be held in the Link 087 (Classroom 3).*

**Course Description:** Introduction to statistics as a science of understanding and analyzing data. Themes include data collection, exploratory analysis, inference, and modeling. Focus on principles underlying quantitative research in social sciences, humanities, and public policy. Research projects teach the process of scientific discovery and synthesis and critical evaluation of research and statistical arguments. Readings give perspective on why in 1950, S. Wilks said, "Statistical thinking will one day be as necessary a qualification for efficient citizenship as the ability to read and write."

*As a quantitative social scientist, I am particularly excited to show you how I use data to answer questions on a variety of topics in political science and public policy. We will also discuss how social scientists and humanists in other fields use data in their research.*

**Course Goals and Objectives:** This course introduces students to the discipline of statistics as a science of understanding and analyzing data. Throughout the semester, students will learn how to make effective use of data in the face of uncertainty: how to collect data, how to analyze data, and how to use data to make inferences and conclusions about real world phenomena.

The course goals are as follows:

1. Recognize the importance of data collection, identify limitations in data collection methods, and determine how they affect the scope of inference.

2. Use statistical software to summarize data numerically and visually, and to perform data analysis.

3. Have a conceptual understanding of the unified nature of statistical inference.

4. Apply estimation and testing methods to analyze single variables or the relationship between two variables in order to understand natural phenomena and make data-based decisions.

5. Model numerical response variables using a single or multiple explanatory variables.

6. Interpret results correctly, effectively, and in context without relying on statistical jargon.

7. Critique data-based claims and evaluate data-based decisions.

8. Complete a research project demonstrating mastery of statistical data analysis from exploratory analysis to inference to modeling.

**Structure of course:** This course will use elements of flipped classroom learning. Before each class session, there will be some readings and video lectures (see “Required Texts” below) to read and watch that focus on core statistical concepts or applications to social science/humanities research. Class will begin with a brief review of concepts and then focus on working through problems together and in groups, answering questions, and discussing practical applications of the statistical concepts for that day. You should also enroll in a lab session. These sessions, led by graduate TAs partnered with an undergraduate TA, will provide guidance on how to use the statistical software environment *R*. It is important to keep up with course material, but I understand that challenges may arise during the semester, so I urge you to reach out to me if you have any concerns.

**Course grades and assignments:***.*

* **Problem Sets:** 20% of final grade. *You will complete graded problem sets on a semi-weekly basis over the course of the semester. I will also suggest ungraded practice problems with answers at the end of the textbook pdf that will help practice the skills you will use on graded assignments. You are welcome to work with others on the problem sets, but must turn in separate assignments that are in your own words. There will be**7 problem sets and I will drop the lowest one at the end of the semester. All problem sets will be weighted equally.*
* **Labs:** 20% of final grade. *You will complete a graded assignment in RStudio/R Markdown after nearly every lab session (9 labs total). You are welcome to work with others on the labs, but must turn in separate assignments that are in your own words. I will drop the lowest lab grade from your final average in this category. All labs will be weighted equally. A ninth ungraded lab will be provided as a guide.*
* **Take Home Midterm 1:** 12.5% of final grade. *This take-home exam will be released at 10 PM on Gradescope and over email on September 28th and will be due at 11:59 PM on September 30th. You may not collaborate with others on the exam, but have the entire 50 hours to work on it.*
* **Take Home Midterm 2:** 12.5% of final grade. *This take-home exam will be released at 10 PM on October 31st and due at 11:59 PM on November 2nd. You may not collaborate with others on the exam, but have the entire 50 hours to work on it.*
* **Project:** 35% of final grade. *You will complete a group project where you use real data to examine a topic of interest using linear regression analysis. As part of this project, you will work with peers in your lab section. I will post more specific directions for the project in September.*
  + **Proposal:** 10% of final grade. *A proposal will be due on Monday October 11th. Lab on October 7th will be dedicated to allowing you to work with your group and get feedback from your TA.*
  + **Lab Report:** 25% of final grade. *The final draft of your project is due by 11:59 PM Eastern on December 3rd, the last day of classes.*

**Required Texts:**

* OpenIntro Statistics (4th Edition, 2019), by Diez, David, Mine Çetinkaya-Rundel, and Christopher Barr. (ISBN: 1943450072).
  + The textbook is available for free online [here](https://leanpub.com/openintro-statistics). A paper copy is also available from Amazon [here](https://www.amazon.com/OpenIntro-Statistics-Fourth-David-Diez/dp/1943450072).
* Coursera Videos: Recorded lectures related to course content are available from Coursera. These are [available](https://www.coursera.org/programs/duke-university-courses-gp9dy?authProvider=duke) to you at no additional cost as a Duke student. It may ask you to sign up for a class with assignments. You can ignore the assignment reminders, any assignments will come from me and will fit the rubric above.
* Other readings are available on Sakai or online. Please let me know if a link to an online resource does not work.

**Statistical Software:**

This course will make use of the programming language *R* through the integrated development environment (IDE) *RStudio*. I have reserved Docker containers [here](https://vm-manage.oit.duke.edu/containers/rstudio) through OIT where you can access the software. You should also download [*R*](https://www.r-project.org/) and [RStudio](https://www.rstudio.com/) for your own computer for free as a backup. Your TA team can help with this in your first lab. You should do assignments in R in the Docker and use R on your computer as a backup.

**Statistics 101 Course Policies**

**COVID and Attendance:** Student health, safety, and well-being are the university’s top priorities. To help ensure your well-being and the well-being of those around you, please do not come to class if you have symptoms related to COVID-19, have had a known exposure to COVID-19, or have tested positive for COVID-19. If any of these situations apply to you, you must follow university guidance related to the ongoing COVID-19 pandemic and current health and safety protocols. If you are experiencing any COVID-19 symptoms, contact student health. 919-681-9355. To keep the university community as safe and healthy as possible, you will be expected to follow these guidelines. Please reach out to me and your academic dean as soon as possible if you need to quarantine or isolate so that we can discuss arrangements for your continued participation in class.

We still find ourselves in challenging times with the pandemic. While you should attend lecture if you are feeling well, if you are feeling ill, please do not attend class. Your absence can be excused through a STINF. Either I or a TA can meet with you over Zoom to discuss whatever you missed in class. If you miss class and would like to catch up on material please either attend office hours (if you are sick, attend virtual office hours), or email me about setting up a time to me so that I or a TA can meet with you.

As of now, I do not plan to formally take attendance in the main section; in the lab, the TAs will take roll, but this will mainly be for the purpose of making sure that I can check in with you to make sure that everything is okay if you are missing classes frequently.

**Other COVID-related policies:** Per Duke policies, you should wear a face mask at all times during class. Please do not eat or drink during class. I understand that this is a long class; if you need a sip of water, please step out of the classroom and then return.

**COVID and Class Flexibility:** Duke is planning to have classes in-person this semester. Given the trajectory of the pandemic however, there is always the possibility of changes needing to be made during the semester. I will let you know of any changes as quickly as I can over email should they need to be made. While I plan to follow the list of assignments and readings/videos below, a change in the semester could cause a need for changes to be made. I again will communicate these as quickly as possible.

**Submission of Assignments**: You will submit all written assignments through either Sakai or *Gradescope*. You will submit labs and the assignments for the project on Sakai, while you should complete the exams and problem sets through *Gradescope*.

**Contact Policy and Office Hours:** Students are encouraged to attend office hours or contact me via email with any questions about the course. I may also available by appointment during the week if one is unable to meet during the times listed on the syllabus. Please contact me over email to see if there might be another mutually workable time in which we can Zoom.

I will hold on-campus office hours outside near Old Chem on Mondays and Wednesday from 1:30 to 2:20. If you and someone else you are working with have the same question, I encourage you to come in groups of up to 2 students. If you have a group of more than 2 students, please pick 2 students from the group to come as representatives of the group. The tables are relatively small and we should try to avoid crowding here.

Since we have office hours right after class, if you have a question about a specific problem, I invite you to come to office hours. I am willing to answer shorter questions as time permits after class and will try to set aside 5-10 minutes at the end of class to answer questions. I ask that rather that crowding the front of the room after class that you remain in your seat and raise your hand if you have a question.

I will also hold Zoom office hours from 5-6:20 on Tuesdays by appointment. You can sign up for these office hours on Calendly here: <https://calendly.com/jacobfhsmith>**.** You may sign up for up to two blocks on a day. You are welcome to sign up in a group of any size for these office hours.

Please feel free to contact me over email as well. I will respond promptly and will do my best to respond the same day to emails; however, I may not always be able to provide an immediate response to a late evening email sent the day before an assignment is due. I will generally plan to check email twice a day during the week and once each day over the weekend. If you have not received a response within 24 weekday hours, please feel free to send me a reminder email.

If you have a broad question about course content, you should begin by posting on Sakai Forums.

TAs will also hold office hours and are available over email. If you have a question in which you would like a quick answer, often the fastest response will be received by going to the office hours of whichever member of the teaching team is holding office hours that day or by posting on Sakai Forums.

If you are going to a Zoom office hour for a TA, when you arrive in the main office hours Zoom room for the member of the teaching team you wish to speak to please put your name on this [Google doc](https://docs.google.com/spreadsheets/d/1Wldj1LTd7bRYjfI-Jv6jIQqoH1tb_imyIipMCeVBOFI/edit#gid=0), indicate which problem/item you have a question about, and please wait for your turn to be invited into a breakout room. Students with questions on the same problem may be invited in together. If you would like to talk one-on-one about something specific to you, please write “individual question” or email to set up an appointment.

**Diversity and Mutual Respect:** It is my intent that students from all diverse backgrounds and perspectives be well-served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that the students bring to this class be viewed as a resource, strength and benefit. It is my intent to present materials and activities that are respectful of diversity and in alignment with [Duke's Commitment to Diversity and Inclusion](https://provost.duke.edu/initiatives/commitment-to-diversity-and-inclusion). Your suggestions are encouraged and appreciated. Please let me know ways to improve the effectiveness of the course for you personally, or for other students or student groups.

Furthermore, I would like to create a learning environment for my students that supports a diversity of thoughts, perspectives and experiences, and honors your identities. To help accomplish this:

* If you feel like your performance in the class is being impacted by your experiences outside of class, please don't hesitate to come and talk with me. If you prefer to speak with someone outside of the course, your academic dean is an excellent resource.
* I (like many people) am still in the process of learning about diverse perspectives and identities. If something was said in class (by anyone) that made you feel uncomfortable, please talk to me about it.

**Late Work:** Deadlines are given in Eastern Time (so Duke time). Please try to get in assignments in by the proper time, but I will be flexible. If you think you may need an extension on any assignment, please do not hesitate to reach out to me directly. This has always been my policy, but it is especially critical during the current crisis. This is a stressful time; I do not want this class to add to your stress unnecessarily. The reason for an extension does not have to be a COVID-related; I will not ask you to explain or to present any documents. Please do try to let me know as far ahead as possible, but I recognize that this will not always be possible.

Generally, unless I have approved a deadline extension in advance, assignments are due according to the deadlines specified, including date and time. Late penalties begin 1 minute after an assignment is due. If an assignment is due at 5 PM Wednesday, the late penalty from 5:01 PM Wednesday through 5 PM Thursday is 5 percentage points (e.g., a 100 would become a 95, a 90 would become an 85). An assignment submitted between 5:01 PM Thursday to 5 PM Friday would lose 10 percentage points. A late lab or problem set will not be accepted more than four days (i.e., 96 hours) late without a valid excuse. If you are submitting a lab after the assignment tab on Sakai has closed, you should submit it to your lab TA team over email and cc me on the email. If you are submitting a problem set after the assignment on Gradescope has closed, you should email it to TA Gaurav Sirdeshmukh, who will upload and grade it. Please also cc: me on this email.

All take-home exams should be submitted by the end of the 48-hour window unless you have spoken in advance with the instructor about other arrangements. All work is due by 11:59 PM on December 3rd unless a valid excuse is approved by the instructor.

**Integrity:** Duke University is a community dedicated to scholarship, leadership, and service and to the principles of honesty, fairness, respect, and accountability. Citizens of this community commit to reflect upon and uphold these principles in all academic and non-academic endeavors, and to protect and promote a culture of integrity. Cheating on exams and quizzes, plagiarism on homework assignments and projects, lying about an illness or absence and other forms of academic dishonesty are a breach of trust with classmates and faculty, violate the [Duke Community Standard](https://studentaffairs.duke.edu/conduct/about-us/duke-community-standard), and will not be tolerated. Such incidences will result in a 0 grade for all parties involved as well as being reported to the [Office of Student Conduct](https://studentaffairs.duke.edu/conduct). Additionally, there may be penalties to your final class grade. Please review [Duke's Academic Dishonesty policies.](https://studentaffairs.duke.edu/conduct/z-policies/academic-dishonesty)

**Grades:**

Letter grades map onto the following scale, which I will use in calculating your final grade for the course (I will round grades in between these ranges up to the nearest whole number):

A+ (97–100), A (93–96), A- (90–92)

B+ (87–89), B (83–86), B- (80–82)

C+ (77–79), C (73–76), C- (70–72)

D+ (67–69), D (63–66), D- (60–62)

F (0-59)

I take grading seriously and am happy to discuss why you received a grade on any assignment during office hours or an appointment at a mutually workable time. Regrade requests must be made within two class days of when the assignment is returned, and must be submitted in writing or using the regrade request in *Gradescope*. These will be honored if points were tallied incorrectly, or if you believe your answer is correct but it was marked wrong. No regrade will be made to alter the number of points deducted for a mistake.

You should not submit a regrade to dispute the number of points deducted for an incorrect response. Please note that by submitting a regrade request, your entire assignment may be regraded and you may potentially lose points. Regrade request will be addressed by the Head TA first or the instructor. No changes on individual assignments will be made after the end of the semester. The only grades changes that will be made will be if the incorrect grade was entered into DukeHub or if the grade was calculated incorrectly (i.e., an incorrect percentage was entered or if the grade was weighted incorrectly).

**Statistics 101 Support & Resources**

**Emergency Notification:** In an emergency, there are several ways that the University will contact you. Campus emergency procedures are described here: <http://emergency.duke.edu>.

**Course Accessibility:** Students who may need special accommodations in this class are encouraged to contact the Student Disability Access Office (<http://www.access.duke.edu/students/index.php>) as soon as possible to ensure that I can implement such accommodations in a timely fashion. I also encourage you to speak to me individually if you have trouble with this process, as I am glad to help with the process if I can.

**Counseling and Psychological Services (CAPS):** Each of you will face some level of challenge during your time at Duke – whether it be a challenge like procrastination, or a more profound challenge that impairs your ability to function. The CAPS staff includes psychologists, clinical social workers, and psychiatrists experienced in working with college-age adults. Information about their services and workshops is available here: <http://studentaffairs.duke.edu/caps/about-us>.

**The Academic Resource Center:** The ARC provides academic support and programming for all Duke undergraduates. Their services include one-on-one consultations and peer tutoring, and they work alongside the Student Disability Access Office to serve students with diagnosed learning disabilities. Their programs include opportunities for students to study together in structured groups ("learning communities"), as well as workshops offered throughout the semester. Further information and resources are available on their website. <http://duke.edu/arc/index.php>.

**The Center for Sexual and Gender Diversity:** This center provides education, advocacy, support, mentoring, academic engagement, and space for LGBTQIA+ and allied students, staff, and faculty at Duke. The Center for Sexual and Gender Diversity also serves and supports Duke alumni/ae and the greater LGBTQ+ community. Further information and resources are available on their website. <https://studentaffairs.duke.edu/csgd>.

*This syllabus has been adapted from previous Statistics 101, Statistics 210, and Writing 101 syllabi. Particularly, I would like to recognize Kelly Moran, Victoria Ellison, Maria Tackett, Márcia Rego, Denise Comer, and Simon Hoellerbauer. The syllabus may be added or modified during the semester and students will be notified in advance of any such change taking effect.*

**Schedule of Readings and Assignments**

**Monday August 23: Data Collection & Observational Studies**

**Watch:**

* The [Introduction video](https://www.coursera.org/learn/probability-intro/lecture/UbGdZ/introduction). (~3 mins.)
* The [Data Basics video](https://www.coursera.org/learn/probability-intro/lecture/Q0zu3/data-basics). (~5 mins.)
* The [Observational Studies & Experiments video](https://www.coursera.org/learn/probability-intro/lecture/Qw8iF/observational-studies-experiments).(~5 mins.)

**Read:**

* OIS Chapter 1, Sections 1.1, 1.2, and start of 1.3 (pp. 9-23; stop after reading section 1.3.2).
* Smith, Jacob. 2016. “Cherie Berry put her picture in every North Carolina elevator. Here’s how that affected her reelection.” *Washington Post Monkey Cage Blog*. <https://www.washingtonpost.com/news/monkey-cage/wp/2016/04/14/cherie-berry-put-her-picture-in-every-north-carolina-elevator-heres-how-that-affected-her-reelection/> (Available on Sakai).

*Please read syllabus before the first day of class.*

**Wednesday August 25: Sampling**

**Watch:**

* The [Sampling and sources of bias video](https://www.coursera.org/learn/probability-intro/lecture/Y96uT/sampling-and-sources-of-bias) (~8 mins.)
* The [Experimental design video](https://www.coursera.org/learn/probability-intro/lecture/Jyc3t/experimental-design) (~3 mins.)
* The [Random Sample Assignment video](https://www.coursera.org/learn/probability-intro/lecture/N6gE8/spotlight-random-sample-assignment) (~4 mins.)
* PBS Clip: [What Went Wrong with Polling in 2020](https://www.pbs.org/newshour/show/what-went-wrong-with-polling-in-2020) (~8 mins.)

**Read:**

* OIS Chapter 1, Sections 1.3 & 1.4 (pp. 22-38).
* Cassino, Dan. 2016. “How Today’s Political Polling Works.” *Harvard Business Review.* <https://hbr.org/2016/08/how-todays-political-polling-works> (Available on Sakai).

***Lab 1 due on Gradescope by 11:59 PM on Monday August 30.***

**Monday August 30: Exploratory Data Analysis**

**Watch:**

* The [Visualizing Numerical Data video](https://www.coursera.org/learn/probability-intro/lecture/9kRJf/visualizing-numerical-data) (10 mins.)
* The [Measures of Center video](https://www.coursera.org/learn/probability-intro/lecture/AM0o6/measures-of-center) (4 mins.)
* The [Measures of Spread video](https://www.coursera.org/learn/probability-intro/lecture/t9Wbk/measures-of-spread) (6 mins.)
* The [Robust Statistics video](https://www.coursera.org/learn/probability-intro/lecture/ssktR/robust-statistics) (1 min.)
* The [Transforming Data video](https://www.coursera.org/learn/probability-intro/lecture/eQa2U/transforming-data) (3 mins.)

**Read:**

* OIS Chapter 2, Section 2.1 (pp. 41-61).
* Ha, Robbie, Peilin La, and Alejandro Ortega. (Faculty Lead Astrid Giugni, Project Manager Jessica Hines, Ph.D.) 2017. *Visualizing Suffering: Tracking Photojournalism and the Syrian Refugee Crisis.* [Watch the Room 351 Video](https://bigdata.duke.edu/projects/visualizing-suffering-tracking-photojournalism-and-syrian-refugee-crisis) and [Read the Executive Summary](https://bigdata.duke.edu/sites/bigdata.duke.edu/files/site-images/Team7ExecSummSlides.pdf).

**Wednesday September 1: More EDA- Intro to Statistical Inference**

**Watch:**

* View the [Exploring Categorical Variables video](https://www.coursera.org/learn/probability-intro/lecture/vEjt0/exploring-categorical-variables) (8 mins.)
* View the [Introduction to Inference video](https://www.coursera.org/learn/probability-intro/lecture/e6IzY/introduction-to-inference) (12 mins.)

**Read:**

* OIS Chapter 2, Sections 2.2 & 2.3 (pp. 61-76).
* Bunyasi, Tehama Lopez and Candis Watts Smith. 2019. *Stay Woke: A People’s Guide to Making All Black Lives Matter.* Chapter 1, pages 29-40. (Available on Sakai, please read sections labeled, “Education,” “Employment,” and “Income and Wealth.”)

***Problem Set 1 Due on Gradescope by 11:59 PM on Wednesday September 1.***

***Lab 2 due on Gradescope by 11:59 PM on Monday September 6.***

***Please fill out Qualtrics survey by noon on Monday September 6. (Will send out on Friday.)***

**Wednesday September 6: Intro to ggplot and dplyr**

**Go through the following tutorials:**

* [Into the Tidyverse](http://maths4.stat.duke.edu:3939/into-tidyverse/)
* [Basics of ggplot](http://maths4.stat.duke.edu:3939/ggplot2-basics/#section-welcome-to-ggplot2)
* [Tidy data wrangling with dplyr](http://maths4.stat.duke.edu:3939/data-wrangling/)

**Wednesday September 8: Probability and Conditional Probability**

**Watch:**

* The [Introduction video](https://www.coursera.org/learn/probability-intro/lecture/07vL4/introduction) (5 mins.)
* The [Disjoint Events + General Addition Rule video](https://www.coursera.org/learn/probability-intro/lecture/qaYwc/disjoint-events-general-addition-rule) (9 mins.)
* The [Independence video](https://www.coursera.org/learn/probability-intro/lecture/D1m0l/independence) (9 mins.)
* The [Probability Examples video](https://www.coursera.org/learn/probability-intro/lecture/GqcO0/probability-examples) (9 mins.)
* The [Disjoint vs. Independent video](https://www.coursera.org/learn/probability-intro/lecture/PSXBC/spotlight-disjoint-vs-independent) (2 mins.)

**Read:**

* OIS Chapter 3, Section 3.1 (pp. 81-94).

***Problem Set 2 Due on Gradescope by 11:59 PM on Wednesday September 8.***

***Lab 3 Due on Gradescope by 11:59 PM on Monday September 13.***

**Monday September 13: Bayes Theorem/Bayesian Inference**

**Watch:**

* The [Conditional Probability video](https://www.coursera.org/learn/probability-intro/lecture/fN6Kh/conditional-probability) (12 min.)
* The [Probability Trees video](https://www.coursera.org/learn/probability-intro/lecture/QE3kX/probability-trees) (10 min.)
* The [Bayesian Inference video](https://www.coursera.org/learn/probability-intro/lecture/qYbZQ/bayesian-inference) (14 min.)
* The [Examples of Bayesian Inference video](https://www.coursera.org/learn/probability-intro/lecture/AWtbY/examples-of-bayesian-inference) (7 min.)

**Read:**

* OIS Chapter 3, Section 3.2 (pp. 95-112)
* Lee, Mike and Benedict King. *The Conversation.* “Bayes’ Theorem: the maths tool we probably use every day, but what is it?” (Available online at <https://theconversation.com/bayes-theorem-the-maths-tool-we-probably-use-every-day-but-what-is-it-76140>.)

**Wednesday September 15: Normal and Binomial Distributions**

**Watch:**

* The [Normal Distribution video](https://www.coursera.org/learn/probability-intro/lecture/M71Nv/normal-distribution) (17 mins.)
* The [Evaluating the Normal Distribution video](https://www.coursera.org/learn/probability-intro/lecture/snzoA/evaluating-the-normal-distribution) (2 mins.)
* The [Working with the Normal Distribution video](https://www.coursera.org/learn/probability-intro/lecture/mwnU6/working-with-the-normal-distribution) (5 mins.)
* The [Binomial Distribution video](https://www.coursera.org/learn/probability-intro/lecture/kJ07f/binomial-distribution) (17 mins.)
* The [Normal Approximation to Binomial video](https://www.coursera.org/learn/probability-intro/lecture/2sosk/normal-approximation-to-binomial) (14 mins.)
* The [Working with the Binomial Distribution video](https://www.coursera.org/learn/probability-intro/lecture/yef4b/working-with-the-binomial-distribution) (9 mins.)

**Read:**

* OIS Chapter, Section 4.1 and start section 4.3 (pp. 133-143 & 149-157).
* Clayton, Aubrey. 2020. “How Eugenics Shaped Statistics.” *Nautilus.* <https://nautil.us/issue/92/frontiers/how-eugenics-shaped-statistics?fbclid=IwAR3ASxPaNGmci_TMIbPQPpAy3NnrxRVoEhpyn5kSjmFnyDg8FVErXDFu82o>.

***Homework 3 Due on Gradescope by 11:59 PM on Wednesday September 15.***

***Lab 4 Due on Gradescope by 11:59 PM on Monday September 15.***

**Monday September 20: Variability in estimates and CLT**

**Watch:**

* The [Introduction video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/EXe3o/introduction) (4 mins.)
* The [Introduction to Inference for Categorical Variables video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/ruAZK/introduction) (3 mins.)
* The [Sampling Variability and CLT for Proportions video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/oJxbr/sampling-variability-and-clt-for-proportions) (15 mins.)

**Read:**

* OIS Chapter 5, Section 5.1 (pp. 170-180).

**Wednesday September 22: Confidence Intervals**

**Watch:**

* The [Confidence Interval for a Proportion video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/nK388/confidence-interval-for-a-proportion) (9 mins.)
* View the [Accuracy vs. Precision video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/QTajX/accuracy-vs-precision) (7 mins.)
* View the [Required Sample Size for a ME video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/QPEXw/required-sample-size-for-me) (4 mins.)
* View the [CI (for a mean) examples video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/txXVo/ci-for-the-mean-examples) (5 mins.)

Read:

* OIS Chapter 5, Section 5.2 (pp. 181-188)
* Mercer, Andrew. 2016. “5 key things to know about the margin of error in election polls.” Pew Research Center. Available at <https://www.pewresearch.org/fact-tank/2016/09/08/understanding-the-margin-of-error-in-election-polls/>.

***Problem Set 4 Due on Gradescope by 11:59 PM on Wednesday September 22.***

***Lab 5 Due on Gradescope by 11:59 PM on Monday September 27.***

**Monday September 27: Midterm 1 Review**

*No additional reading- study for exam.*

*I will be available during normal class time to answer questions.*

**Wednesday September 29: No Class- Midterm I**

***Midterm exam released at 10 PM on September 28th due at 11:59 PM on September 30.***

***No lab on Thursday September 30.***

**Monday October 4: Fall Break**

**Wednesday October 6: Hypothesis Tests**

**Watch:**

* The [Hypothesis Test for a Proportion video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/VAo5K/hypothesis-test-for-a-proportion) (9 mins.)

**Read:**

* OIS Chapter 5, Section 5.3 (pp. 189-201).
* Denworth, Lydia. 2019. “The Significant Problem of P-Values.” *Scientific American.* (Available on Sakai.)

***Lab on October 7 will focus on group work on proposal for projects. Proposal due by 11:59 PM on Monday October 11.***

**Monday October 11: Inference for a Single Proportions**

**Read:**

* OIS Chapter 6, Section 6.1 (pp. 208-216).

**Wednesday October 13: Inference for Comparing Two Proportions**

**Watch:**

* The [Estimating the Difference Between Two Proportions video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/kI4Ma/estimating-the-difference-between-two-proportions) (17 mins.)
* The [Hypothesis Test for Comparing Two Proportions video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/vMH7n/hypothesis-test-for-comparing-two-proportions) (13 mins.)
* The [Small Sample Proportions video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/B7mb4/small-sample-proportions) (10 mins.)
* The [Examples video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/w7VQF/examples) (4 mins.)
* The [Comparing Two Small Sample Proportions video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/rUhQw/comparing-two-small-sample-proportions) (5 mins.)

**Read:**

* OIS, Chapter 6, Section 6.2 (pp. 217-228).

***Problem Set 5 Due on Gradescope by 11:59 PM on Wednesday October 13.***

***Lab 6 Due on Gradescope by 11:59 PM on Monday October 18.***

**Monday October 18: Inference with t-distribution**

**Watch:**

* The [Introduction to Inference for Numerical Variables video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/xtUR0/introduction) (4 mins.)
* Please review The [Confidence Interval (for a mean) video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/DA30M/confidence-interval-for-a-mean) (11 mins.) If you watched previously, okay to skip/skim.
* The [Sampling Variability and CLT video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/lkQnZ/sampling-variability-and-clt) (20 mins; okay to skip around, some of this is similar to what we discussed for the CLT for proportions)
* The [t-distribution video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/FlRrd/t-distribution) (7 mins.)
* The [Inference for a mean video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/qs7Ml/inference-for-a-mean) (9 mins.)

**Read:**

* OIS Chapter 7, Sections 7.1 and 7.2 (pp. 251-266).
* Clarke, Andrew. 2010. *The Blue Dog Coalition: Impact of a Single Issue Caucus from the 104th to 110th Congress.* Read Chapter 1, pp. 1-12. (Available on Sakai, pages labeled at bottom right of page.)

**Wednesday October 20: Inference for comparing two means**

**Watch:**

* The [Inference for comparing two independent means video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/wkwlZ/inference-for-comparing-two-independent-means) (8 mins.)
* The [Inference for comparing two paired means video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/k5zhM/inference-for-comparing-two-paired-means) (9 mins.)

**Read:**

* OIS, Chapter 7, Sections 7.4 & 7.5 (pp. 267-284).
* Clarke, Andrew. 2010. *The Blue Dog Coalition: Impact of a Single Issue Caucus from the 104th to 110th Congress.* Read Chapter 3, pp. 39-60, section labeled “The Source of Preference.” (Available on Sakai, pages labeled at bottom right of page.)

***Problem Set 6 Due on Gradescope by 11:59 PM on Wednesday October 20.***

***Lab 7 Due on Gradescope by 11:59 PM on Monday October 25.***

**Monday October 25: Chi-Square and ANOVA**

**Watch:**

* The [Chi-Square GOF Test video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/OO6iS/chi-square-gof-test) (14 mins.)
* The [Chi-Square Independence Test video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/LEIm3/the-chi-square-independence-test) (11 mins.)
* The [Comparing more than two means video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/TI3DD/comparing-more-than-two-means) (6 mins.)
* The [ANOVA video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/KoTvZ/anova) (9 mins.)
* The [Conditions for ANOVA video](https://www.coursera.org/learn/inferential-statistics-intro/lecture/hSgp3/conditions-for-anova) (2 mins.)

**Read:**

* OIS Chapter 6, Section 6.3 (pp. 229-239) and 7.5 (285-298).
* Clarke, Andrew. 2010. *The Blue Dog Coalition: Impact of a Single Issue Caucus from the 104th to 110th Congress.* Read Chapter 3, pages 61-75, section labeled “Representation on Prestige Committees.” (Available on Sakai, pages labeled at bottom right of page.)

**Wednesday October 27: Midterm 2 Review**

**No lab on Thursday October 28- use the time to review for the second exam.**

**Monday November 1: Midterm Exam 2**

**Midterm exam released at 10 PM on October 31st and due at 11:59 PM on November 2nd.**

**Wednesday November 3: Intro to Bivariate Regression**

**Watch:**

* The [Introduction video](https://www.coursera.org/learn/linear-regression-model/lecture/UcKYt/introduction) (~1 min.)
* The [Correlation video](https://www.coursera.org/learn/linear-regression-model/lecture/QP6Mw/correlation) (~9 mins.)
* The [Residuals video](https://www.coursera.org/learn/linear-regression-model/lecture/FpKWn/residuals) (~2 mins.)
* The Least Squares Line video (~12 mins.)
* The [Conditions for Linear Regression video](https://www.coursera.org/learn/linear-regression-model/lecture/MKPvv/conditions-for-linear-regression) (~10 mins.)

**Read:**

* OIS Chapter 8, Sections 8.1 and start of 8.2 (pp. 305-321).

***Lab 8 Due on Gradescope by 11:59 PM on Monday November 8.***

**Monday November 8: Outliers and Inference for Bivariate Regression**

**Watch:**

* The [Prediction and Extrapolation video](https://www.coursera.org/learn/linear-regression-model/lecture/aHBzs/prediction-and-extrapolation) (~4 mins.)
* The [R2 Video](https://www.coursera.org/learn/linear-regression-model/lecture/lMej8/r-squared) (~4 mins.)
* The [Regression for Categorical Explanatory Variables Video](https://www.coursera.org/learn/linear-regression-model/lecture/FdxeN/regression-with-categorical-explanatory-variables) (6 mins.)
* The [Outliers in Regression](https://www.coursera.org/learn/linear-regression-model/lecture/10xAG/outliers-in-regression) Video (~7 mins.)
* The [Inference in Linear Regression](https://www.coursera.org/learn/linear-regression-model/lecture/icLpS/inference-for-linear-regression) Video (~12 mins.)

**Read:**

* OIS, Chapter 8, Finish Section 8.2 and Sections 8.3 and 8.4 (pp. 322-337).

**Wednesday November 10: Multiple Regression I- Inference**

**Watch:**

* The [Inference for Multiple Linear Regression](https://www.coursera.org/learn/linear-regression-model/lecture/Pg7rK/inference-for-mlr) Video (~12 mins.)

**Read:**

* OIS, Chapter 9, Section 9.1 (pp. 343-352).
* Abramowitz, Alan. 2016. “Will Time for Change Mean Time For Trump?” *PS: Political Science and Politics* 49(4): 659-660. (Available on Sakai; will update this for a 2020 article if Abramowitz writes one.)
* Clarke, Andrew. 2010. *The Blue Dog Coalition: Impact of a Single Issue Caucus from the 104th to 110th Congress.* Read Chapter 3, pages 76-80, section labeled “Blue Dog Association with Fiscal Conservatism” (Available on Sakai, pages labeled at bottom right of page.)
* Abramowitz, Alan. 2020. “It’s the Pandemic, Stupid! A Simplified Model for Forecasting the 2020 Presidential Election.” *Sabato’s Crystal Ball* (Available online at <http://centerforpolitics.org/crystalball/articles/its-the-pandemic-stupid-a-simplified-model-for-forecasting-the-2020-presidential-election/>.)

***Problem Set 7 Due on Gradescope by 11:59 PM on Wednesday November 10.***

***Lab 9 Due on Gradescope by 11:59 PM on November 15.***

**Monday November 15: Writing and Statistics**

**Read:**

* Masket, Seth. 2011. “The Perils of Holding a Tea Party at High Altitude: Colorado’s Senate and Gubernatorial Races in 2010.” In *Pendulum Swing*, Edited by Larry J. Sabato. Boston: Longman. (Available on Sakai.)
* Canes-Wrone, Brandice, David Brady, and John Cogan. 2002. “Out of Step, Out of Office: Electoral Accountability and House Members' Voting.” *American Political Science Review* 96(1): 127-40. (Available at on Sakai).

**Wednesday November 17: Multiple Regression II- Model Selection and Diagnostics**

**Watch:**

* The [Model Selection](https://www.coursera.org/learn/linear-regression-model/lecture/pR8Mi/model-selection) Video (~11 mins.)
* The[Diagnostics for Multiple Linear Regression](https://www.coursera.org/learn/linear-regression-model/lecture/quxtC/diagnostics-for-mlr) Video (~8 mins.)

**Read:**

* OIS Chapter 9, Sections 9.2 & 9.3 (pp. 353-362).

**Lab on November 18 is a project work period.**

**Monday November 22: Other Types of Regression**

**Read:**

* Smith, Jacob. 2020. “Just as Electable: Black Democratic Candidates in Swing Districts.” *Politics, Groups, and Identities.* Available as an Online First Articles. (Available on Sakai.)
* Smith, Jacob. 2020. “Black candidates can win in swing districts.” (Available at <https://theconversation.com/black-candidates-can-win-in-swing-districts-151980>.)
* Benjamin, Andrea, Ray Block Jr., Jared Clemons, Chryl Laird, and Julian Wamble. 2020. “Set in Stone? Predicting Confederate Monument Removal.” PS: Political Science and Politics 53(2): 659-660. 237-242.

***Class on November 24 and lab on November 25 cancelled for Thanksgiving holiday.***

**Monday November 29/ Wednesday December 1: Project Work Day/Extra Catch up Day if we fall behind-TBD later in the semester.**

**Lab on Thursday December 2 is a project work day.**

**Project due December 3 at 11:59 PM.**