SOC 301-01 - Social Statistics Syllabus

Pacific University - Spring 2017

Dr. Chester Ismay

- **Course meeting time**: Tuesdays and Thursdays 2:45 PM - 4:20 PM in Price Hall, Room 214
- **Course webpage**: http://ismayc.github.io/soc301-f2017
- **Office Hours**
  - Tuesdays and Thursdays 1:00 PM - 2:30 PM in Scott Hall 100B5 in the Center for Languages & International Collaboration (CLIC)
  - By appointment
- **Office Phone Number**: Not available (Happy to meet in the Library/CLIC or another place on campus as needed though.)
- **Email**: chester@pacificu.edu
- **Slack**: @chester (Preferred and **highly encouraged** way of contact)

Tutors

We have three excellent tutors/TAs for this course. Their office hours and where they will be located will be announced soon. Dinisa will frequently join us during our class meeting. A photo and more information about them is below.

- Dinisa Malcom
- Kawita Kaur
- Griffin Fraser

Overview

Welcome to the Age of Data, where information is all around us, helping us live happier, healthier lives. Or does it? Do we know yet if cell phones cause cancer? Have we come to a decision on whether we should be eating lots of meat or none at all to stay healthy? Despite all of this information, it can be challenging to turn it into the knowledge from which we can make sound decisions. Statistics is the field that aims to bridge this gap between information and knowledge and this course is an application-oriented introduction to modern statistical modeling and inference in the social sciences.

The main focus of this course will be on data visualization as the bridge to statistical and data science topics from a social perspective. We will avoid the traditional, formula-based methodologies of "plug-and-chug" statistics courses and instead use the statistical computing language R to help build intuitions and solve problems in a reproducible way. We will strive to identify the correct problem types and focus on the interpretation of results and plots.

Course Goals

Upon leaving this course you'll be able to answer the following questions:

- What is a p-value and how is it used in social science literature? How does it relate to a confidence interval and why are both of them important topics for the social sciences?
- What makes an effective plot for a given set of variables?
- What goes into making a good statistical analysis? What assumptions are made?
- How do we start a statistical analysis? What questions should we ask?

- How can we create an effective story with data?

We will be using DataCamp, which provides an interactive way for you to learn how to use R to conduct analyses and make plots, throughout the course. You can expect very frequent assignments to be due using it. The expectation is that by the end of the course you are able to conduct statistical analyses and produce appropriate plots and do so in a reproducible environment. We will also be shifting much of this work in R to R Markdown and using an RStudio Server as the course progresses. You'll find that R is the best way to provide customizability and reproducibility and I believe is the best way to use the tools that statisticians, data scientists, and scientific researchers are using today without being too overwhelming.

Student Learning Objectives

Upon completion of this course, students who have been full participants in the course should be able to:

- Identify tidy data sets by carefully specifying the observational units and the types of variables that make up the data set.
- Make and interpret the Five Named Graphs: scatter-plots, line-graphs, histograms, boxplots, and barplots.
- Use multivariate thinking by producing and discussing appropriate plots showing the relationships between three (or more) variables.
- Identify the appropriate plot and type of analysis needed to answer a given social research question.
- Describe the process of simulation to build inferential statistics concepts.
- Explain the results of hypothesis testing and confidence intervals in layman's terms.
- Check that assumptions are met in order to perform traditional statistical tests (t-tests, ANOVA, regression, etc.) and be able to provide an alternative randomization-or-bootstrap-type analysis if assumptions are not met.
- Utilize correlation and regression to begin the process of statistical modeling.
- Use R, RStudio, and R Markdown to conduct reproducible, well-documented, statistical analyses.
- Discuss how social science research depends on carefully made statistical graphics and analyses.

Course Description from the Catalog

Introductory statistics course for students in the social sciences. The emphasis of the course will be on understanding how social scientists use numerical data to understand social phenomena, and how to use and interpret statistical measures and techniques commonly reported in the social sciences literature. Prerequisite: MATH 122. Does not meet Social Sciences core requirement (2010 catalog). 4 credits.

Textbook

A MODERN DIVE into Data with R (2016) by Chester Ismay and Albert Y. Kim. The online HTML version is preferred but a PDF version is also available. This textbook is free and open-source and will be a key component of the course. You'll frequently be asked to read through the textbook and complete the Learning Checks and Review Questions for practice prior to class to check for comprehension.

The textbook is more general in focus than social statistics specifically, but it is intended to be as user-friendly and engaging as possible. The assignments throughout the course will be more applied to social statistics in particular.

Additional Resources

Getting used to R, RStudio, and R Markdown (2016) by Chester Ismay. This is also a free online resource I put together to help completely new users of R not be quite as intimidated in learning what might initially be scary. We'll refer to this document frequently throughout the course after the first few weeks of work at DataCamp.com.

OpenIntro: Introductory Statistics with Randomization and Simulation (2014), by David M. Diez, Christopher D. Barr, and Mine Çetinkaya-Rundel, available in three formats: pdf, tablet-friendly pdf, and paperback edition. You'll notice a pattern here: the textbook is also free and open-source.

Class components

This course has four components: Problem Sets, DataCamp Assignments, Group Project, and Exams/Quizzes.

Problem Sets

Problem sets reinforce the concepts from the reading as well as the material that we address in class. You're encouraged to work on the problem sets with fellow students both outside and during class. We will sometimes work through the problem sets together in class, but you should expect to be asked to hand in your problem sets at the end of class each time they are assigned. Late problems sets will **not** be accepted, so you are encouraged to work on them before and/or during class meetings. If you do work on the assignment with other students (no more than 2), please turn in the assignment as a group and clearly note who you worked with at the top of your sheet of paper.

We will also have frequent checking for understanding using Plickers and other low stakes quizzing strategies in class. This is done to ensure the coverage of topics is at the right level while allowing you the chance to check your own understanding as much as possible in a low stress environment.

Extremely important note

In order to pass the class (with a grade of C- or better), you **MUST** complete AT LEAST 75% of the Problem Sets and DataCamp Assignments. Students that fall behind in the class are the ones that feel like the course didn't meet their expectations in my experience. I will support you, but if you aren't doing the work for the course, you will not receive an A, B, or C grade in the course. If you feel like you are falling behind, you are encouraged to schedule a time to meet with me (on a weekly basis if needed).

DataCamp Assignments

Almost every week you'll be working with data using R and completing assignments on our DataCamp course page. These assignments will usually be due on Tuesdays at 1:00 PM Pacific time and Thursdays at 1:00 PM Pacific time. DataCamp allows me to track when and how often you are working on the assignments and your progress, so if you'd like to make a good impression and boost your learning, I encourage you to start these assignments as soon as possible and ask me questions as needed. The best way to ask questions is in person during my office hours or in meeting with the tutors. If this doesn't work for you, you should take a screenshot of your webpage showing the error and what you've tried as clearly as possible. You can then include the image to the #general discussion on Slack or as a private message to me there.

Group Project

You will also have the opportunity to use RStudio Server to complete a group project with your peers near the middle of the semester. This group project will include your writing of a data journalism piece in the style of http://FiveThirtyEight.com. It will include carefully created graphics and clean descriptions to a layperson and will serve as a culminating experience of the first half of the course. The Group Project will likely be assigned on or before March 2nd, a draft of your Group Project will be due on March 14th, and your final revised Group Project will be due on March 23rd at 10 AM Pacific time.

Exams/Quizzes

We'll have one midterm exam and several quizzes throughout the semester in order to challenge your understanding and provide me and you with a sense of where you're at. Each of these assessments will build on previous knowledge and you can think of them as being **cumulative** in much the same way the final exam will be. This is to your benefit. I want to help you learn how to study cumulatively and practice so that final exams aren't as terrifying. I'll provide plenty of tips and we'll practice a lot in class to help you with this. My goal is for you to see how the course all fits together whenever possible.

Quizzes will be multiple choice and the questions will frequently come from a subset of questions that you will each create and turn in during the class period before the quiz is to be given. Quizzes will first be taken individually and then immediately following in groups of about three. Exams will be more traditional, pencil-and-paper assessments, but you will be allowed to use RStudio cheatsheets on the exams. More details on this as the semester progresses.

A cumulative final exam will be given on May 15th from noon to 2:30 PM.

Grading Criteria

This course will provide you with many opportunities to demonstrate your learning of statistical concepts and your ability to apply them to a variety of social data sets and contexts. DataCamp Assignments will encourage and allow you to continue to attempt the problems until you have obtained the right answer and will give you immediate feedback on whether your response is correct or not. I'll also provide you with opportunities to show me that you have gained knowledge of concepts later on in the course if you initially struggled.

That being said, I do EXPECT hard work and focus from you as students. I want you to aspire to learn and grow as students and individuals. The content from a course that is too easy is quickly forgotten at the end of the semester and a course that is too difficult can cause anxiety that interferes with genuine learning opportunities. I strive to find the right balance of challenge and fairness for you.

Your course grade will be determined by the following basis with relative weights given:

- Attendance/Participation/Problem Sets – 5%
- DataCamp Assignments - 10%
- Group Project – 15%
- Midterm Exam - 20%
- Cumulative Quizzes – 25%
- Final Exam - 25%

Note the policy mentioned above about 75% of Problem Sets and DataCamp Assignments being completed to receive a grade higher than a D in the course.

I won't assign or give extra credit in this course. There are plenty of opportunities to show me (and yourself) that you are improving and learning the material. Getting behind will get you in trouble quickly though as I've found getting out of a rut academically is far harder than you may think.

Grading System

I strongly encourage you to focus on your own learning of the materials and how they align with the goals for the course I have laid out. You should identify for yourself how well you are understanding the material and be able to assess this frequently. This will allow you to keep the knowledge that you obtain during the course much after the semester is over. I am required to assign you letter grades though and they will be assigned as follows:

A: 90-100%; B: 80-89%; C: 70-79%; D: 60-69%; F: less than 60%

These cut-offs may be lowered at the end of the semester and I reserve the right to give +/- grades as I see fit.

Course Schedule

The course schedule is available in the link at the top of this webpage.

What to Expect in Class

As much as possible, I will devote the first five to ten minutes of class and the last five minutes of class to activities designed to help you retain the material you have just encountered or to reflect on previous content. It is imperative that you value these opportunities and your attendance and participation during them will help you succeed in the course.

We'll be working through problems together in class and discussing content from your reading assignments. This class may be different from other classes you take or your experiences in classrooms from the past. I prefer to think of myself as your guide on the side instead of the sage on the stage. I'm here to assist you as needed, but I don't want to be the passer of all knowledge to you. In other words, I'm hoping to create an environment where you develop understanding based on each of your own needs rather than just dictating how and what to think to you.

You'll be actively working to learn about social statistics inside the classroom. I will lecture from time to time, but this will not be a class where you can passively sit back and copy everything from the board. You'll need to play along with R, to answer and ask questions during class, to take careful and thoughtful notes, and to work well with your classmates to solve problems.

Most importantly, I want the classes to be fun and I want you to value what we cover. It's my job to make it interesting; it's your job to put in the effort and have a growth mindset about statistics. You'll be amazed at the end of the semester just how much knowledge you've gained and how your hard work has paid off.

Course attendance policy

Attendance will be collected and active participation in in-class and out-of-class activities is expected. **THREE** absences of any sort are allowed. That being said I understand all the demands that are on you as students. Keep me posted on what is going on if you must miss class. These three absences can serve any purpose you like.

If you have a medical emergency or another valid institutional reason for missing class, we can set up a plan together. You'll find that missing classes hurts your understanding of the material and map of how the whole course fits together. Each additional class missed beyond the third absense will result in a **10%** grade reduction on your final course grade. I take TIMELY attendance very seriously.

Makeup quizzes

Makeup quizzes and exams are rarely given. It isn't fair to the other students in the class to take an assessment at a later date and it hurts your flow in the class because you'll be playing catch-up with other material.

What to Expect Outside of Class

This course **requires** you to read the textbook and to check for understanding as you learn new material and review past material. You'll be working on Learning Checks and Review Questions from the textbook and doing statistical analysis and making graphics using R. This will be time-consuming, especially if you aren't focusing on them, but it is my goal to separate these assignments into smaller pieces. This will help ensure that you aren't procrastinating until the last minute on a big assignment and also that you are using the new knowledge gained as much as possible just when you learn it.

Tips for success:

1. Spend 15 minutes every day at least five days a week quizzing yourself on concepts from the course. I recommend you create flash cards as the course progresses and ask yourself to provide the definition when given a term or to explain what particular R code produces. The process of forcing yourself to think about the concept even if you can't get the answer exactly right will help you learn.

2. Don't just read over your notes or the book multiple times. One of the most common things I hear from students that didn't perform well on a quiz or assignment is that they studied the material a lot by reading over the book many times. This doesn't help you learn the content. It creates a false sense of understanding by creating an association with the words. YOU HAVE TO QUIZ YOURSELF if you want to learn and to succeed.

3. Create your own quizzes. One of the best strategies you can do if you have had trouble with exams and quizzes in the past is to identify the key ideas from the class and create your own problems to solve in an environment similar to the one in which you take quizzes and exams. Try as much as you can to create problems that you can relate to that involve the content. This can be your everyday life, your specific major, or any other interest you may have. Real-life connections help to build actual connections in your brain, thus, helping you learn.

4. Make mind maps that relate all of the content throughout the course together. By creating a diagram that connects concepts together you'll see how everything is related and that will help you create the story of the course instead of seeing concepts as disjoint pieces.

5. Study with someone else, especially with the R material. You'll need to practice with R nearly every day if you want to learn how to use it. It's going to be like learning a foreign language at the beginning. You'll need to learn the syntax and practice "speaking it" as much as possible. You'll likely catch on quickly, but you won't if you don't practice often.

6. Ask questions of me, of the course tutors, and of your classmates. I'm always open to answering questions and you should never feel like your question isn't appropriate. You should also look for opportunities to help others and teach them the content whether it be in person or online using Slack. This will help you as a learner as well.

7. Enjoy the learning process! Failure to not understand everything all at once is to be expected. Our brains need to struggle to learn. You can't build muscles by sitting on the couch and you can't learn statistics by not struggling a little bit. Devote some time every day to thinking about a more difficult problem than what is given to you. How would you solve it? What's more difficult about it? Evaluate your learning and find out what works best for you.

Additional Information

Academic Misconduct Policy

Pacific University has no tolerance for academic misconduct/cheating. It is university policy that all acts of misconduct and dishonesty be reported to the Associate Dean for Student Academic Affairs. Sanctions that may be imposed for such misconduct range from an "F" for the assignment, an "F" for the course, and suspension or dismissal from the university. Forms of academic misconduct include but are not limited to plagiarism, fabrication, cheating, tampering with grades, forging signatures, and using electronic information resources in violation of acceptable use policies.

Learning Support Services for Students with Disabilities

If you have documented challenges that will impede your learning in any way, please contact our LSS office in Clark Hall (ext.2717; lss@pacificu.edu). LSS staff will meet with students, review the documentation of their disabilities, and discuss the services that Pacific offers and any appropriate ADA accommodations for specific courses. Please follow up with me about how we can best match your accommodations if you need them.

Tutoring and Learning Center (TLC)

The TLC is located in Scott Hall, 1st-floor. The center focuses on delivering one-on-one and group tutoring services for foreign languages, math and science courses, and writing skills in all subjects. Students should consult with the center's director for information on tutoring available for other subjects. Day and evening hours; walk-ins welcome!

Unauthorized Recordings

Students are prohibited from making audio and/or visual recordings of lectures or presentations without prior consent of the instructor or presenter.