about this course // syllabus

See the schedule for a list of topics by week, with links to slides.

Description of the course

This graduate course aims to provide students with practical understanding of and experience with core concepts and methods in modern data analysis. The focus is on biological data, but skills will be transferable to other disciplines. Students will become familiar with major topics in univariate and multivariate statistics, analysis of large data sets, and Bayesian analysis. There is a particular emphasis on modeling and conceptual understanding of statistical noise and uncertainty. The course is advanced in that we will move through the material quickly with the goal of providing a solid foundation for subsequent learning. Students will learn to use the powerful statistical programming language R, and the flexible modeling package Stan.

Instructors:

- Peter Ralph, main instructor: plr@uoregon.edu
  – office hours: Tuesday 3-4:30pm in 275 Onyx or by appointment
- Ananya Kapoor, graduate teaching assistant: akapoor@uoregon.edu
  – office hours: Wednesday 10am-noon (471 MCK)
- Bill Cresko, supporting instructor: wcresko@uoregon.edu

Course Information

- **Schedule:** Tue/Thur 10:00-11:50 AM in B040 PSC and on zoom (see Canvas for link)
- **Websites:** besides the page you are looking at, assignments and announce-
  ments will be distributed on the Bi 610 canvas page
- **Assignments:** Assigned, and due, on Thursdays.

Textbooks

We will be assigning reading from these books:

- Quinn, G. & M. Keough. 2002. *Experimental Design and Data Analysis
  for Biologists.* Cambridge Univ. Press.

There is a list of other useful reading materials on the page of references.

**Software:**

- Latest version of R (install here)
- A working version of RStan
- An R scripting environment such as RStudio,
  - or a good code editor like Atom or BBEdit or vim
  - *optional:* Git (here is an intro)

**Prerequisites:**

Students should be comfortable with algebra and the basics of calculus, have some familiarity of introductory statistics, and have some experience with programming.

**Class structure, remote learning, and support**

I aim to make class available remotely, by broadcasting lectures on zoom (see Canvas for the link) and making the recordings available afterwards on Canvas. However, for most students in-class participation is more engaging and results in better learning outcomes, so I strongly encourage in-person attendance. However, I know that things happen in life that might make you unable to come – for instance, please do not come to class if you are sick. (Also note that UO currently requires students and employees to be fully vaccinated.)

If you are unable to turn in assignments on time, please let me know when you are able, so I can make alternative arrangements.

**Assignments and evaluation**

Course grades will be based entirely on assignments. Most assignments are weekly homeworks, assigned on Thursdays and due the following Thursday. These will be short, readable, descriptive statistical reports, which you will write in Rmarkdown (here are some tips for doing this), and so also contain the R code that produces the analyses and figures. Once or twice during the quarter instead of a report we will also assign peer reviews, where you review a classmate’s assignment (here’s how to do this). There will also be a few other types of assignment: for instance, filling out a class survey (graded on participation only), and (in Winter quarter) a group data analysis project. The contribution of these to the final grade will be proportional to the points assigned, and thresholds for letter grades will be chosen to reflect a graduate-level course: A: good, demonstrates competence in all topics and mastery of many; B: acceptable, demonstrates competence in all topics; C: unsatisfactory.
Inclusion and accessibility

Please tell us your pronouns and/or name, especially if it differs from the class roster. We take seriously our responsibility to create inclusive learning environments. Please notify us if there are aspects of the instruction or design of this course that result in barriers to your participation! You are also encouraged to contact the Accessible Education Center in 164 Oregon Hall at 541-346-1155 or uoaec@uoregon.edu.

We are committed to making our classroom an inclusive and respectful learning space. Being respectful includes using pronouns for your classmates. Your classmates come from a diverse set of backgrounds and experiences; please avoid assumptions or stereotypes, and aim for inclusivity. Let us know if there are classroom dynamics that impede your (or someone else’s) full engagement.

Please see this page for more information on campus resources, academic integrity, discrimination, and harassment (and reporting of it).
course schedule

The (Rmarkdown) source code for these lectures is available at the github repository, or by replacing the `.slides.html` suffix with `.Rmd` in the link below; the slides are made using reveal.js. Here are the slides from Fall 2021/Winter 2022, Fall 2020/Winter 2021, Fall 2019/Winter 2020, and Fall 2018/Winter 2019. Unlike previous years, this year (2022), there will only be a Fall term.

**Fall 2022**

**Week 1 (9/27)** Overview of data science - description and estimation, uncertainty and simulation, hypothesis testing.
- Slides: Introduction
- how to git the slide source
- Reading: Quinn & Keough chapters 1-4
- Slides: Hypothesis testing and p-values
- Slides: The t distribution
- Slides: Confidence intervals
- Homework 1 *(due 10/6)*

**Week 2 (10/4)** Visualization, confidence intervals, permutation tests, and the bootstrap.
- Slides: Visualization
- Slides: The bootstrap
- Slides: The central limit theorem
- Slides: Permutation tests
- Reading: Quinn & Keough chapters 1-4 (still)
- Homework 2 *(due 10/13)*

**Week 3 (10/11)** Linear models, ANOVA, and formulas.
- Reading: Quinn & Keough chapter 5, 6, 8

**Week 4 (10/18)** Multiple testing, error rates, and some history.
- Reading: Quinn & Keough chapter 7

**Week 5 (10/25)** Random effects and mixed models; intro to Bayesian stats.
• Reading: Quinn & Keough chapter 13
• Reading: Kruschke, chapters 1, 2

**Week 6 (11/1)** Bayesian model fitting, Markov chain Monte Carlo
• Reading: Kruschke, chapters 4, 5, 6, 7

**Week 7 (11/8)** Logistic models and GLMs; sharing power
• Reading: Quinn & Keough chapter 13
• Reading: Kruschke, chapters 9, 10

**Week 8 (11/15)** Robust linear models; data analysis example
• Reading: Donihue, C.M., Herrel, A., Fabre, A.C. et al. *Hurricane-induced selection on the morphology of an island lizard.*

**Week 9 (11/22, no class Thursday)** Factor analysis, dimensionality reduction, and PCA
• Reading: Quinn & Keough chapter 17

**Week 10 (11/30)** Factor analysis continued, t-SNE; random forests
• Reading: The two cultures, by Leo Breiman
Campus resources and policies

Campus resources to support your learning

Tutoring and Academic Engagement Center: Drop-in math and writing support in addition to tutoring, study skills support, and Class Encore. Located in the 4th Floor Knight Library. (541) 346-3226, engage@uoregon.edu

Counseling Center: Call anytime to speak with a therapist who can provide support and connect you with resources. Located on the 2nd Floor of the Health Center. (541) 346-3227.

Accessible Education Center: If there are aspects of the instruction or design of this course that result in barriers to your participation, please notify us as soon as possible. You are also encouraged to contact the Accessible Education Center. If you are not a student with a documented disability, but you would like to let the Center know about class issues that will impact your ability to learn, we encourage you to come visit the Center during our office hours so that we can strategize how you can get the most out of this course. Located on the 1st Floor of Oregon Hall (541) 346-1155, uoaec@uoregon.edu

Center for Multicultural Academic Excellence (CMAE): has a mission to promote student retention and persistence for historically underrepresented and underserved populations: “We develop and implement programs and services that support retention, academic excellence, and success at the UO and beyond. We reaffirm our commitment to all students, including undocumented and tuition equity students.” Located on the 1st Floor of Oregon Hall (541) 346-3479, cmae@uoregon.edu

The UO Access Shuttle is an on-campus ride service provided at no cost to students with conditions that limit mobility. More information and a sign-up form can be found at: parking.uoregon.edu/content/access-shuttle.

Duck Rides is a service that provides free, inclusive, and accessible alternatives to traveling alone at night for UO students, faculty, and staff.

Academic integrity

All students are expected to adhere to the University’s guidelines on academic integrity as outlined in the Student Conduct Code: policies.uoregon.edu/vol-3-
administration-student-affairs/ch-1-conduct/student-conduct-code. Academic misconduct includes cheating (“any act of deception by which a student misrepresents or misleadingly demonstrates that the student has mastered information on an academic exercise that the student has not mastered”), and plagiarism (“using the ideas or writings of another as one’s own”). All persons involved in academic dishonesty will be disciplined in accordance with University regulations and procedures.

**Discrimination and Harassment**

**Prohibited Discrimination and Harassment:** Any student who has experienced sexual assault, relationship violence, sex or gender-based bullying, stalking, and/or sexual harassment may seek resources and help at safe.uoregon.edu. To get help by phone, a student can also call either the UO’s 24-hour hotline at 541-346-7244 (SAFE), or the non-confidential Title IX Coordinator at 541-346-8136. From the SAFE website, students may also connect to Callisto, a confidential, third-party reporting site that is not a part of the university.

Students experiencing any other form of discrimination or harassment can find information at respect.uoregon.edu or aaeo.uoregon.edu or contact the non-confidential AAEO office at 541-346-3123 or the Dean of Students Office at 541-346-3216 for help. As UO policy has different reporting requirements based on the nature of the reported harassment or discrimination, additional information about reporting requirements for discrimination or harassment unrelated to sexual assault, relationship violence, sex or gender based bullying, stalking, and/or sexual harassment is available in the Employee Responsibilities section of the Office of Investigations and Civil Rights Compliance website.

**Reporting:** The instructor of this class is a Student-Directed Employee. If you disclose discrimination or harassment to me, I will respond to you with respect and kindness. I will listen to you, and will be sensitive to your needs and requests. I will not judge you. I will support you. I will only report the information shared to the university administration when you as the student requests that the information be reported (unless someone is in imminent risk of serious harm or is a minor). Please note the difference between ‘privacy’ and ‘confidentiality.’ As a Student-Directed Employee I can offer privacy because I am not required to report certain information to the university. However, I cannot be bound by confidentiality in the same way that a counselor or attorney is. Confidential resources such as these means that information shared is protected by federal and state laws. Any information that I as a student-directed employee receive may still be accessed by university or court proceedings. This means, for example, that I could still be called as a witness or required to turn over any related documents or notes that I keep.

Please note also that I am required to report all other forms of prohibited discrimination or harassment to the university administration. Specific details about confidentiality of information and reporting obligations of employees can
be found at titleix.uoregon.edu.

**GE involvement and conflict of interest**

This course has a graduate student GE, who will be involved in evaluating (grading homeworks) and instructing (office hours). If you have any privacy, conflict of interest, or other concerns about their role, then please contact us (the instructors) and we will address them (for instance, by grading your work ourselves).