EDS 222: Statistics for Environmental Data Science

Quarter: Fall 2021
Units: 4
Grading: Letter
Prerequisites: None required

Instructor information, meeting times, & materials

Instructor: Tamma Carleton
Email: tcarleton@ucsb.edu
Office hours: Tuesday 3pm - 4pm PST (Location: 3418 Bren Hall, pending COVID conditions)
The best way to contact me is by email (I will do my best to respond within 24 hours)

Class meets:
  Lecture: T 9:30am - 10:15am PST (Bren Hall 1414)
  Lab: Th 9:30am - 10:15am PST (Bren Hall 1414)

Course website: https://tcarleton.github.io/EDS-222-stats/

Required textbook(s) / reader(s):
• *Introduction to Modern Statistics*, by Mine Çetinkaya-Rundel and Johanna Hardin [link]
• [Optional] *Geographic Information Analysis*, by David O'Sullivan and David Unwin [link]

Computing requirements:
• Minimum MEDS device requirements
• R version 4.0.2 (or higher)
• RStudio version 1.4.1103 (or higher)
Basic course information

Statistics is the science of collecting, manipulating, and analyzing empirical data. In this class, we will learn the statistical fundamentals that will enable us to draw conclusions about the environment and its interaction with social and economic systems. We will cover fundamental statistical concepts and tools, and then apply and expand upon those tools to learn some temporal and spatial statistical methods that are particularly helpful in environmental data science. Welcome!

**Learning objectives:** The goal of this course is to enable MEDS students to confidently and competently apply statistical tools to environmental and socio-environmental datasets.

- Master fundamental statistical concepts and learn to apply them in R
- Design, conduct, and interpret linear and nonlinear regression models
- Apply spatial statistical methods to environmental data
- Learn about and address common spatial and temporal dependency issues that arise in environmental datasets

Course details

**Course components:** All lecture slides and lab materials will be posted to the course website. Homework assignments will be posted, submitted, and graded on our Github Classrooms website.

**Weekly learning schedule:**
- Tuesday lecture: 75 minutes (slides & whiteboard)
- Thursday labs: 75 minutes (code-along lessons & group problem-solving)
- Homework: (Mostly) weekly, (mostly) posted and due on Thursdays at 5pm

**Important dates:**
- Midterm exam: October 28th (in-class)
- No class (holidays): November 11th & 25th
- Final project write-up due: December 2nd
**Reading:** In order to make time for hands-on learning during Thursday labs, we will move quickly to cover content in Tuesday lectures. Students are strongly encouraged to read posted readings *before* Tuesday lectures so that content is familiar.

**Assignments:** Homework assignments will be critical to meeting learning objectives in this course. Assignments will ask you to take the content from lectures and apply it to new datasets and contexts, building on applications completed in Thursday Lab sessions. In some cases, assignments will directly expand on problems completed in Labs, but in other cases will entail new data and new types of problems. Assignments will generally be assigned weekly (posted and due Thursdays at 5pm), although their pace will slow later in the course when students spend more time working on final projects. You are permitted to work with at most two other people on each assignment but must list your collaborators’ names at the top of your write-up. *Each assignment will be graded with probability 0.5.* If an assignment is not graded, full marks are given to everyone who submits a complete assignment.

**Final project:** During the second half of the course, students will develop a final project that applies the tools learned in this course to a new question, dataset, and context of their choosing. The details of the project will be finalized later in the course, but it will likely have both a write-up and a presentation component. The final project will be conducted individually.

**Grades:** The course requirements include homework assignments (~6 assignments, 40%), a midterm exam (25%), a final project (25%), and class participation (10%). The midterm will be in-class and closed book.

**Getting extra help:** Come to office hours! We don’t have a TA this first year of MEDS, so please communicate with me and/or MEDS Staff if you’re having a hard time keeping up and we will do what we can to get you the support you need.

**Slack:** We will have an #eds-222 channel in the MEDS workspace. Please feel free to use this as a space to communicate with each other about the content of the course. I will monitor the Slack channel irregularly and send relevant links and other information, but do not expect me to respond rapidly to Slack messages regarding questions on course content. Please use office hours or email for direct communication about course content.
**Tentative course schedule and topics:** The first half of this course is dedicated to statistical fundamentals. After the midterm, we begin learning spatial and temporal statistical tools that are particularly relevant to environmental data science work. This schedule is subject to change; please check the [course website](#) for updates.

<table>
<thead>
<tr>
<th>Week</th>
<th>Lecture</th>
<th>Lab</th>
<th>Reading</th>
</tr>
</thead>
</table>
| 1    | No class | Course introduction and motivation; Sampling from a population; | • Chapter 2 in *IMS*  
• *Rolf, et al. (2021)*  
[optional] Esther Duflo video on RCTs |
| 2    | Categorical vs numerical data; summary statistics; law of large numbers; probability density functions; Bayes Rule; skewness; outliers; relationships between variables | Visualizations and summaries of categorical data in R; simulations in R; summary statistics and measures of skew in R | • Chapters 4-6 in *IMS* |
| 3    | Correlation; conditional expectations; ordinary least squares; measures of model fit | Correlations; visualizations of relationships between variables in R; OLS in R | • Chapters 7-8 in *IMS* |
| 4    | Logistic regression; interactions and categorical variables in linear regression | OLS in R, continued | • Chapters 8-10 in *IMS* |
| 5    | The basics of inference; the Central Limit Theorem; hypothesis testing; confidence intervals | Inference and hypothesis testing in linear regression | • Chapters 13; 16-22 in *IMS* |
| 6    | Inference in linear regression, cont’d | Midterm Exam | • Chapters 24-26 in *IMS* |
| 7    | Time series analysis; serial correlation; panel data basics | Time series in R; temporal lag models | • TBD |
| 8    | Point processes; clustering algorithms | No class | • TBD |
| 9    | Vector fields; interpolation; kriging; spatial weighting | Implementing kriging, weighting, and interpolation algorithms in R | • TBD |
| 10   | Spatial autocorrelation; moving window statistics and regression; spatial lag models | No class |   |
| 11   | TBD/slack time | Final project presentations |   |
| 12   | Final project presentations |   |   |
Access, accommodations and conduct

**Course conduct:** All students are expected to read and comply with the [UCSB Code of Conduct](#). We are committed to making this course a welcoming and inclusive environment for everyone, regardless of gender, gender identity and expression, race, age, sexual orientation, disability, physical appearance, or religion (or lack thereof). We expect all students to help ensure this environment of inclusivity and will not tolerate harassment of any form.

**Access and Accommodations:** Please submit requests for accommodations often and early. It is never too late to apply for DSP accommodations. If you have any kind of disability, whether apparent or non-apparent, learning, emotional, physical, or cognitive you may be eligible to use formal accessibility services on campus. To arrange class-related accommodations, please [contact DSP](#). DSP will initiate communication about accommodations with faculty. By making a plan through DSP, appropriate accommodations can be implemented without disclosing your specific condition or diagnosis to course instructors.

**Additional student resources**

*The text below is provided by the UCSB Disabled Students Program.*

**Counseling and Psychological Services (CAPS).** As a student you may experience a range of issues that can cause barriers to learning, such as strained relationships, increased anxiety, alcohol/drug problems, feeling down, difficulty concentrating and/or lack of motivation. These mental health concerns or stressful events may lead to diminished academic performance or reduce your ability to participate in daily activities. CAPS is available to assist you with addressing these and other concerns you may be experiencing. You can learn more about the broad range of confidential mental health services available on campus. They can be reached by phone at 805.893.4411, or online at [http://caps.sa.ucsb.edu](http://caps.sa.ucsb.edu). The CAPS building is the pink building next to the Humanities and Social Science building (HSSB)

**Food insecurity:** [http://food.ucsb.edu/](http://food.ucsb.edu/) includes the Cal Fresh Program [http://food.ucsb.edu/calfresh](http://food.ucsb.edu/calfresh) and the Associated Students food bank [https://foodbank.as.ucsb.edu](https://foodbank.as.ucsb.edu)
Resource Center for Sexual and Gender Diversity (RCSGD) in the SRB, offers a host of services for LGBTQI+ students including a library and many events throughout the year. [http://rcsgd.sa.ucsb.edu/](http://rcsgd.sa.ucsb.edu/)

Dream Scholars/Undocumented Student Services Program offers workshops, helps students find scholarships and financial support as well as providing community for our undocumented students. [http://www.sa.ucsb.edu/dreamscholars/home](http://www.sa.ucsb.edu/dreamscholars/home)

**Campus Learning Assistance Services (CLAS)** helps students grow academically by offering workshops, walk-in and pre-scheduled tutoring, and writing help both for native and non-native (ESL) English as a second language speakers. Over 50% of students will stop by CLAS at one time or another. [http://clas.sa.ucsb.edu](http://clas.sa.ucsb.edu)

**Student Resource Building (SRB)** houses many campus resources offices, including the African Diasporic Cultural resource Center, the American Indian Resource Center, the Asian Resource Center, the Middle Eastern Resource Center, the Non-Traditional and Re-Entry Student Resource Center. [http://www.sa.ucsb.edu/student-resource-building/home](http://www.sa.ucsb.edu/student-resource-building/home)

**Multicultural Center (MCC)**, located in UCEN, hosts a wide variety of cultural events and educational programming throughout the year, including film showings, lectures, musical performances, and more: [http://mcc.sa.ucsb.edu/](http://mcc.sa.ucsb.edu/)

**Campus Advocacy, Resources, & Education (CARE)** offers 24/7 confidential support and advocacy in situations of sexual assault, dating and domestic violence, and stalking. Located in the SRB, they can be reached at 805.893.4613 or [http://wgse.sa.ucsb.edu/care/home](http://wgse.sa.ucsb.edu/care/home)

**Financial Crisis Response Team**: If you are experiencing issues of housing insecurity contact the Financial Crisis Response Team at financialcrisis@sa.ucsb.edu to begin application for assistance.

**Health and Wellness**: Student well-being is integral to academic success, student development, and life satisfaction. On this website, students will find links to a range of services related to well-being such as: assistance with basic needs (food, housing, finances); counseling and physical health resources, daily wellness centers and programs; social connection, and personal safety. [https://wellbeing.ucsb.edu/](https://wellbeing.ucsb.edu/)