

Syllabus

ES 207 Environmental Data Analysis

University of California Merced

Spring 2018

Prof. Erin L. Hestir

12/20/2017

- General Course Information
 - Instructor
 - Meetings
 - Format
- Course Description
 - Course Goals
 - Learning Outcomes
- Course Assignments and Grading
 - Homework (50%)
 - Literature Assignments (25%)
 - Literature Assignment: Oral Presentation (10%)
 - Literature Assignment: Written Review (10%)
 - Literature Assignment: Reading Discussion Participation (5%)
 - Mid-Term Exam (10%)
 - Final Exam (15%)
- Required Reading
- Computing Requirements
 - Hardware
 - Software
 - Operating System
 - Analysis Software
 - Online Resources
- Course Policies
 - Accommodations for Students with Disabilities:
- Notes on this class:
- Tentative Schedule
- Some Useful Campus Resources
 - Counseling and Psychological Services
 - Discrimination & Sexual Violence Prevention
 - CARE Office
 - Food Assistancess (HEROES)

General Course Information

Instructor

Prof. Erin L. Hestir elhestir@ucmerced.edu (<mailto:elhestir@ucmerced.edu>)

- If you email me please place ES 207 in the beginning of the subject line

Office Hours: Tuesdays 12:00-1:00 pm or by appointment (SE1 212)

Meetings

Monday & Wednesday 3:00-4:15 pm CLSSRM 262

Format

Lectures are on Wednesdays, Labs are on Mondays. There will be an assigned reading due before lecture or lab each day. Please come to class prepared to discuss the reading. We will have a reading discussion EVERY day.

Course Description

The objective of this class is to provide students with probabilistic and statistical methods to analyze environmental data. This class emphasizes both theoretical and applied aspects of data analysis methods. Weekly lab exercises are from environmental applications. Topics include: distribution, hypothesis test, linear regression, multiple regression, uncertainty analysis, outlier detection, sample design, and spatial and temporal data analysis.

Course Goals

This course will provide students with a hands-on overview of the major data analysis techniques, contemporary problems in and common approaches to environmental sciences. It will provide students with hands-on experience using contemporary software to conduct environmental data analysis, challenge students to use best practices in data management, emphasize the importance of reproducibility in research, and foster interpersonal communication and collaboration.

Learning Outcomes

Upon successful completion of this course, students will be able to:

1. Identify problems in environmental science requiring data collection, analysis, and dissemination
2. Successfully collect and analyze, and disseminate results
3. Communicate the theory, basis, methods, results, and context of applied environmental data analysis
4. Will be proficient in R and R Studio
5. Will be familiar with RMarkdown, Shiny, and GitHub

Course Assignments and Grading

Homework (50%)

We will complete eight homework assignments designed to give you familiarity with different data analysis techniques and proficiency with R. No late assignments will be accepted; however, the lowest grade will be dropped. Several of the assignments build off of previous assignments. So, if you don't complete one homework assignment, you will have twice as much work to do for the next one. You will be required to format all of your homework in RMarkdown, including fully documented scripts. All homework assignments will be turned in as pdf documents knit through RMarkdown.

Literature Assignments (25%)

- *Oral Presentation*- 10%
- *Written Review*- 10%
- *Discussion Participation*- 5%

You will select a peer reviewed article from the literature of your interest (e.g., hydrology, landscape ecology, meteorology, etc.) that uses a particular analytical technique (e.g., regression, ordination) as assigned by the instructor. You will critically analyze the methods of the study, assess whether they were appropriate in the context of the study, and identify strengths and weaknesses in the approach. You will post the article on the course website by the article selection due date. Everyone will be expected to read all of the papers posted, and will be expected to participate in literature discussions following the presentation schedule.

Literature Assignment: Oral Presentation (10%)

You will lead a discussion of the paper in class, which will include a short oral presentation of the context of the study, explanation of the methods used, and a facilitated discussion of the strengths and weaknesses of the approach.

Literature Assignment: Written Review (10%)

1,000-1,500 words on the strengths and weaknesses of the paper methodology and a discussion of the analytical technique used in the paper (with outside references beyond the paper). You will be assessed on the intellectual content of the review, the theoretical soundness and quality of the critique, as well as the quality, grammar, and style of your writing. This is grad school, writing matters. A lot.

Literature Assignment: Reading Discussion Participation (5%)

Participation is expected in reading discussions

Mid-Term Exam (10%)

Take-home exam, cumulative

Final Exam (15%)

Take-home exam, cumulative

Required Reading

There are two required texts, and several ancillary electronic readings that are also required. Ancillary electronic readings will be posted on the course website

1. A Primer of Ecological Statistics
(<http://www.sinauer.com/media/wysiwyg/tocs/PrimerEcologicalStatistics2.pdf>) 2nd Ed. (Sinauer) by Nicholas Gotelli and Aaron Ellison. *First edition is acceptable.*
 - Available at the campus bookstore, online (e.g., Amazon, Barnes and Noble), or as an ebook through the library
2. R for Data Science (<http://r4ds.had.co.nz/>) (O'Reilly) by Garrett Grolemund and Hadley Wickham.
 - Available **free** online, or available in print from online retailers

Computing Requirements

Hardware

You are expected to bring a laptop to class everyday. We will walk through examples related to the homework, work on key components of the homework, and practice hands-on data analysis.

You can bring any laptop you like, though I recommend the following configuration requirements (anything less may leave you frustrated because processing might go slowly and you may run out of space depending on how much housekeeping you do on your hard drive).

- **Processor**
 - Recommended: Intel Core i5-6xxx or equivalent
 - Minimum: Intel Core i3-3xxx or equivalent
- **Memory**
 - Recommended: 8 GB RAM or more
 - Minimum: 4 GB RAM
- **Storage**
 - Recommended: 500 GB internal hard drive
 - Minimum: 500 GB internal hard drive

Software

Basic computing skills are expected for anyone getting a STEM degree. It is 2018. If you are uncomfortable with computers, now is the time to get comfortable. There is no better way to learn how to use a computer than to use one. Therefore, you are responsible for all installation, troubleshooting, and hardware maintenance. This is part of the learning process and it differs from machine to machine. You get to figure this out for your machine. It is fun! Really.

Operating System

Your choice (mostly)! R works pretty well across many platforms. Although I recommend having the most recent install of whatever OS you prefer.

- Windows Vista/7/8/10
- Mac OSX 10.6+ (64-bit)
- Ubuntu 12.04-15.10/Debian 8 (32-bit)
- Ubuntu 12.04-15.10/Debian 8 (64-bit)
- Ubuntu 16.04+/Debian 9+ (64-bit)
- Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)

- Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)

Analysis Software

You will need to install the following software before the first class meeting. Please make sure you have installed the latest versions of the following open-source software packages as appropriate for your OS:

- QGIS (<https://www.qgis.org/en/site/forusers/download.html>)
- R (<https://cran.r-project.org/>)
- R Studio (<https://www.rstudio.com/products/rstudio/download/#download>)
- R Markdown (<http://rmarkdown.rstudio.com/>)
- Tidyverse (<https://www.tidyverse.org/>)
- Shiny (<https://shiny.rstudio.com/>)
- Git + GitHub (<http://r-pkgs.had.co.nz/git.html>)(make sure to create an account on GitHub using the free plan)

Online Resources

We are working with open source software. That means that it is constantly changing, but also that there is more help for you out on the web than you could possibly imagine! I highly recommend bookmarking the following pages for this class and your life beyond this class:

- <https://www.r-project.org/> (<https://www.r-project.org/>)
- <https://www.r-bloggers.com/> (<https://www.r-bloggers.com/>)
- <https://www.rdocumentation.org/> (<https://www.rdocumentation.org/>)
- <https://stackoverflow.com/questions/tagged/r> (<https://stackoverflow.com/questions/tagged/r>) (my personal favorite)
- <https://www.statmethods.net/> (<https://www.statmethods.net/>) (my other personal favorite, especially when I'm making graphs)
- <https://www.rstudio.com/resources/cheatsheets/> (<https://www.rstudio.com/resources/cheatsheets/>) (Consider printing some or all of these in color for a quick desk reference-that's what I do)
- <https://stats.stackexchange.com/questions/tagged/r> (<https://stats.stackexchange.com/questions/tagged/r>)

ACCORDING TO THE FUNDAMENTAL LAW OF MEDIA GRAPH CONSTRUCTION, ANY TWO THINGS THAT LOOK CORRELATED ON A GRAPH ARE CAUSALLY RELATED IN REAL LIFE.



IT FOLLOWS THAT ANY TWO THINGS CAN BE MADE TO HAVE A CAUSAL RELATION BY HAVING TWO FUNCTIONS ON ONE GRAPH WITH TWO Y-AXES AND SCALING THE Y-AXES UNTIL THEY KINDA LOOK LIKE EACH OTHER.



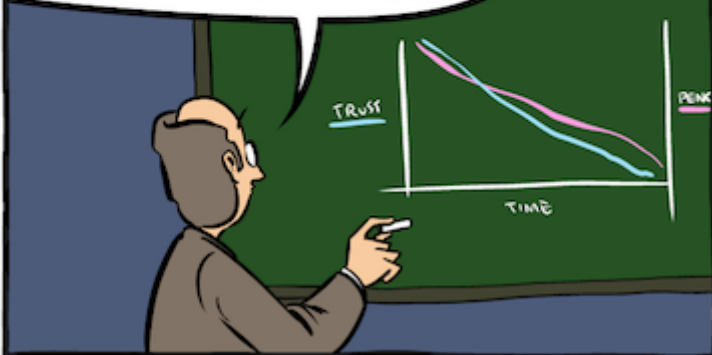
BY THIS MEANS, ANY TWO TRENDS CAN BE FORCED INTO CAUSAL RELATION FOR THE PURPOSES OF AN ARTICLE, BOOK, OR SPEECH.



WHY DO YOU THINK PEOPLE TRUST THE NEWS LESS THAN EVER?



ACCORDING TO THIS GRAPH, PENGUINS ARE AT FAULT.



Zach Weinersmith, Saturday Morning Breakfast Cereal <https://www.smbc-comics.com/?id=3129>
(<https://www.smbc-comics.com/?id=3129>)

Course Policies

- **Attendance & Participation:** Attendance and full participation is expected. You are a valued and critical component of the UC Merced campus and your respective graduate program. You were admitted to your program because we value your intellect, experience and background. Please be prepared to share this with us. This class relies heavily on YOU to work.
- **Academic Integrity:** This is graduate school, and there is absolutely no excuse for anything less than complete compliance with the spirit and letter of academic honesty. Academic honesty is the cornerstone of modern science and intellectual achievement. You should fully embrace this as a budding scientist, and hold yourself and others to the highest standards. Students will be held to the UC Standards of Conduct and the UC Merced Academic Honesty Policy. Presume all work in the class is to be conducted and completed independently unless otherwise instructed for a specific assignment.
- **Behavior:** Don't be a jerk. Your fellow students are your peers not just now, but forever from this point on. The people in this classroom are the beginnings of your budding professional network. One day they will be your collaborators, your bosses, your program managers, your anonymous reviewers. Treat your professional network with care and respect.
- **Late assignments and makeup work:** The work in this class is cumulative, and to fall behind is to fail. No late assignments will be accepted. A partial grade is better than a zero grade.

Accommodations for Students with Disabilities:

The University of California Merced is committed to ensuring equal academic opportunities and inclusion for students with disabilities based on the principles of independent living, accessible universal design and diversity. I am available to discuss appropriate academic accommodations that may be required for student with disabilities. Requests for academic accommodations are to be made during the first three weeks of the semester, except for unusual circumstances. Students are encouraged to register with Disability Services Center to verify their eligibility for appropriate accommodations.

Notes on this class:

1. To make this course accessible to a diverse audience, the lectures assume little-to-no knowledge of the subject matter. However, we ramp up quickly to sophisticated topics and the cumulative nature of the material makes it important that you complete the readings and assignments in a timely manner. This is graduate school: it is a lot of work. I'm going to make you do a lot of work, both reading and analysis-wise. But I expect this course will pay dividends on the back end as you can take the scripts you generate and knowledge you gain with you and apply it to your thesis, dissertation, and other research projects.
2. This class is specifically designed for you. You will get as much out of this as you put into it. Make the most of it. I want to enable you to begin your journey into Environmental Data Analysis and Data Science. It is a long journey, and this is only the beginning.
3. The class is intended to expose you to a range of different ideas and techniques in environmental data analysis. Note, this is only exposure. Given the wide range analytical approaches and methods in the environmental sciences, there is no possible way a single course can cover it all. Our ground rules are:
 - Don't be intimidated.
 - Don't be afraid to ask questions.

- Don't be surprised if I don't know all the answers: I'm still learning too!
4. Just as environmental data science is ever-evolving, so is this syllabus. My goal is to give you useful skills and tools to take away with you, and we will adapt this syllabus to meet those needs together throughout the semester.

Tentative Schedule

This schedule is subject to change

Tentative Course Schedule. Subject to Change.

Date	Lecture.Topic	Lab.Topic	Reading.Discussion	Assessment
17-Jan	Intro to Environmental Data Analysis			
22-Jan		Intro to R/ R Markdown & Github	R4ds Ch. 1: "Introduction"	
24-Jan	Probability, Distributions, & Summary		Gotelli & Ellison Ch. 1-3	
29-Jan		Exploratory data analysis in R & visualization with ggplot2	R4ds Ch. 2: "Explore"	HW 1 due
31-Jan	Statistical Inference Frameworks & Data Science		Gotelli & Ellison Ch. 4-5	
5-Feb		Data wrangling in R: Welcome to the TidyVerse	R4ds Ch. 3: "Wrangle"	HW 2 due
7-Feb	Experimental & Sampling Design		Gotelli & Ellison Ch. 6-7	
12-Feb		R programming: Pipes, functions, data types	R4ds Ch. 4: "Program"	HW 3 due
14-Feb	Data Transformation, Curation, & Management		Gotelli & Ellison Ch. 8	
19-Feb	Presidents Day Holiday			Paper Selection due
21-Feb	Correlation, analysis of variance, comparing means and ranks		Helsel & Hirsch Ch. 5-8	
26-Feb		More with R Markdown, ggplot 2, and an intro to Shiny	R4ds Ch 5: "Communicate"	HW 4 due
28-Feb	Models I: Linear, Quantile, Logistic & Multiple Regression		Gotelli and Ellison Ch. 9	

Date	Lecture.Topic	Lab.Topic	Reading.Discussion	Assessment
5-Mar		To loop or not to loop & Debugging		
7-Mar	Models II: GLM, GAM, RLM			midterm assigned
12-Mar		Model Building in R Part 1	R4ds Ch. 4: "Model"	midterm due
14-Mar	Multivariate Analysis I: Ordination & Classification		Gotelli and Ellison Ch. 12	
19-Mar		Model Building in R Part 2	Literature assignment	
21-Mar	Time Series & Signals		Literature assignment	HW 5 due
26-Mar	Spring Recess			
28-Mar	Spring Recess			
2-Apr		Trend analysis	Helsel and Hirsch Ch. 12	
4-Apr	Spatial Analysis I		Wegmann et al. Ch. 2	
9-Apr		Handling spatial data in R	Literature assignment	HW 6 due
11-Apr	Spatial Analysis II		Literature assignment	
16-Apr		Raster processing and image classification	Literature assignment	
18-Apr	Spatial Analysis III		Literature assignment	
23-Apr		Change detection	Literature assignment	HW 7 due
25-Apr	Spatial Analysis III		Literature assignment	
30-Apr		Species distribution models	Literature assignment	
2-May	Wrap-up			HW 8 due; final assigned

Date	Lecture.Topic	Lab.Topic	Reading.Discussion	Assessment
9-May	Final Exam	8:00-11:00 am		final due



If she loves you more each and every day,
by linear regression she hated you before you met.

Zach Weinersmith, Saturday Morning Breakfast Cereal, <http://www.smbc-comics.com/comic/2011-08-05>
(<http://www.smbc-comics.com/comic/2011-08-05>)

Some Useful Campus Resources

Counseling and Psychological Services

The mission of UC Merced Counseling and Psychological Services (CAPS) (<http://counseling.ucmerced.edu/>) is to support the mental health and well-being of our students. It is the intention of all CAPS staff to provide a safe, confidential atmosphere of acceptance and accessibility to professionals in the field of psychology.

Contact Information (Confidential Help)

Phone: (209) 228-4266

counseling@ucmerced.edu (<mailto:counseling@ucmerced.edu>)

Discrimination & Sexual Violence Prevention

The University of California is committed to creating and maintaining a community where all individuals who participate in university programs and activities can work and learn together in an environment free of harassment, exploitation or intimidation (<http://dsvp.ucmerced.edu/>).

Contact Information

Phone: (209) 285-9510

msalvador2@ucmerced.edu (mailto:msalvador2@ucmerced.edu), Michael Salvador, Director of Compliance,

CARE Office

Campus Advocacy, Resources, & Education (CARE) (<https://care.ucmerced.edu/>) provides prevention education for the UC Merced community to achieve an environment free from the threat of sexual violence, dating/domestic violence, and stalking. They provides free and confidential assistance for all UC Merced affiliates (including Undergraduate students, Graduate students, Staff and Faculty. Stop by KL 107.

Contact Information (Confidential Help)

Campus Advocate: Val (209) 386-2051

Valley Crisis Center 24/7 Hotline (209) 722-4357

Food Assistances (HEROES)

CalFresh (<http://heroes.ucmerced.edu/what-is-Calfresh>) is a monthly stipend system that allows you to purchase food for no cost at all on your part. If you qualify for work study you most likely qualify for CalFresh.

Contact Information

Phone: 209-228-4187

heroes@ucmerced.edu (mailto:heroes@ucmerced.edu)