ESM 244: Advanced Data Analysis

Bren School of Environmental Science & Management (UCSB)

Instructor: Allison Horst (ahorst@bren.ucsb.edu)
Office Hours: TBD
Course materials: allisonhorst.github.io

Summary: ESM 244 is a survey course in advanced topics in statistics and data analysis for environmental scientists (ordinal and multinomial logistic regression, bootstrapping, non-linear models, intro to time-series analysis, spatial data analysis and interpolation, ordination methods, cluster analysis, text mining, etc.) while continuing to build skills and habits in data science (data management, organization, manipulation, analysis, reproducibility, and collaboration).

Thursday lectures will often include coding, so bring your computer to Thursday lectures (as well as labs). Labs and course assignments will be completed using R Markdown in R/RStudio with GitHub. Since the lab will be held in a Bren lecture room (1414), you need to bring your charged laptop (and charger) with R and RStudio installed to each lab.

Term projects (2):
- A personal website/blog featuring at least 4 data science projects, created with blogdown in R
- A Shiny app

Detailed guidelines and expectations for the term projects will be posted during Week 1.

Grading:
- Assignments: 35%
- Week 4 Quiz: 5%
- Term project 1 (data sci projects website): 20%
- Term project 2 (Shiny app): 40%

There is no required reader or textbook for this course. Readings and course materials will be posted on the course website (allisonhorst.github.io).

WEEKLY TOPICS (VERY TENTATIVE):
- Week 1: ESM 206 Review + blogdown introduction
- Week 2: Multiple linear regression recap, PCA & RDA intro, Shiny intro
- Week 3: Binary & ordinal logistic regression intro, Shiny apps continued
- Week 4: Bootstrapping & non-linear models (take-home readings, lab & quiz this week)
- Week 5: Spatial data visualization, wrangling and analysis
- Week 6: Spatial data analysis & interpolation (kriging)
- Week 7: Time series wrangling, data visualization, decomposition, exploration
- Week 8: Time series data continued (intro to forecasting)
- Week 9: Text mining and sentiment analysis
- Week 10: Hierarchical cluster analysis
COURSE POLICIES:

- Assignments submitted late will only be accepted within one week of the due date, and will be worth 50% of the original score. **Homework submitted more than one week after the original due date will not be accepted.**
- Assignments are due at the beginning of lecture on due dates.
- Lab attendance is mandatory
- There will be no make-up assignments/midterms/quizzes or opportunities for extra points or extra credit due to low scores (but please talk with us about your circumstances and if you need support or extra consideration)
- Cheating/plagiarism (including R code) will result in 0 points awarded for the assignment or midterm/quiz/exam and disciplinary action according to UCSB policy
- If you are worried about your overall grade and/or passing the class, please reach out to us to discuss your concerns