ESM 244: Advanced Data Analysis

Bren School of Environmental Science & Management (UCSB)

Instructor: Allison Horst (ahorst@bren.ucsb.edu)
Office: Bren Hall 4406
Office Hours: TBD

Summary: ESM 244 provides an introduction to advanced topics in statistics and data analysis (ordinal and multinomial logistic regression, bootstrapping, non-linear models, intro to time-series analysis, spatial data analysis and interpolation, introduction to Bayesian data analysis, and principle components analysis) while continuing to build skills and habits in data science (data management, organization, manipulation, analysis, reproducibility, and collaboration). The class is a survey course in advanced topics; the structure is to introduce a wide breadth of topics, skills and tools.

Weekly lab attendance is mandatory. Labs and course assignments will be completed using Rmarkdown in R/RStudio with GitHub. Since the lab will be held in a Bren lecture room (1414), you **need to bring your charged laptop with R and RStudio to each lab.**

Assignments will be posted biweekly (4 total). There will be several tutorials posted to GauchoSpace that will not be graded, but should be completed individually. For group assignments, all members of the group are expected to contribute to, and understand, the entire assignment submitted. Assignments may involve oral presentations.

Final Project: Your final project will be to build a Shiny app in R, using data of your choice, which you will present during the last week of class. Detailed expectations and grading criteria will be posted to GauchoSpace. There will be intermediate check-ins throughout the quarter to ensure that you are making progress on your app.

Grading: Assignments (60%), Final (40%)

There is no required reader or textbook for this course. Readings and course materials will be posted on the course GauchoSpace site.

WEEKLY TOPICS (TENTATIVE):

Week 1: ESM 206 Review + Shiny apps introduction
Week 2: Ordinal/multinomial logistic regression
Week 3: PCA/RDA
Week 4: Bootstrapping, nonlinear models
Week 5: Longitudinal data + panel regression intro
Weeks 6 - 7: Time series data visualization, decomposition, analysis and forecasting
Week 8 - 9: Spatial data visualization, wrangling and analysis
Week 10: Cluster analysis, text 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 (9:30am) on due dates. Anything submitted after 9:35am on the due date is considered late and will be worth 50% of the original score. **You should expect Bren printer problems and traffic jams when considering this policy.**
- You are required to wait two full days before asking the instructor or TAs about scores on all returned course material
- 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
- 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, it is **your** responsibility to reach out to us to discuss your concerns