Stat 41: Topics in Statistics: Data Analysis Policy Projects
Fall 2017, Swarthmore College

Class Meetings: Tuesdays and Thursdays, 9:55-11:10, Martin 210

Prerequisites: Stat 11 or Stat 21 or permission of the instructor

Professor:
Lynne Steuerle Schofield
Science Center 148
lschofi1@swarthmore.edu
610-328-7896
Office Hours: Mon 2:30-4:30, Thurs 2:30-4:30 or **by appointment**.

Course description: In this course, students work in teams on a semester-long data analysis problem. A key objective of the course is to expose students to the variety of challenges faced by the data analyst. Students research the scientific background of their problem, consult with area organizations, and communicate their methods and results both in writing and in class presentations. At the end of the semester, each team will present their findings in class to their area organization and write a final report for use by the organization. There will be presentations during the second week of classes describing the projects further.

In Stat 41, we will extend the skills in exploring data and interpreting results that were begun in Stat 11 and Stat 21, by building and fitting models, investigating model assumptions, interpreting results from the statistical models, and report writing. We will do this using real data to solve real problems. We will work concurrently on two tracks: lectures on statistical techniques that may be useful in your projects, and regular project meetings within your groups, with me, and with your area organizations.

Textbooks: Chatterjee and Hadi, *Regression Analysis by Example*, 4ed, Wiley.

Software: We will mainly use the statistical software package R which is available free on-line at http://www.r-project.org. I recommend also downloading RStudio (http://www.rstudio.com/products/RStudio/). My expectation is that you have seen R and RStudio in a previous class. If you have not, please come see me.

Accommodations: If you believe that you need accommodations for a disability, please contact the Office of Student Disability Services (Parrish 113W) to arrange an appointment to discuss your needs (email: studentdisabilityservices@swarthmore.edu). As appropriate, the Office will issue students with documented disabilities a formal Accommodations Letter. Since accommodations require early planning and are not retroactive, please contact the Office of Student Disability Services as soon as possible. Visit the Student Disability Service Website at (http://www.swarthmore.edu/academic-advising-support/welcome-to-student-disability-service) for details about the accommodations process. You are also welcome to contact me privately to discuss your academic needs. However, all disability-related accommodations must be arranged through the Office of Student Disability Services.

Grading: Final grades will be based on a total score, with homework worth 25%, the midterm worth 20%, the final written report worth 25%, oral presentations worth 20%, your group/self/org assessments worth 8% and class participation worth 2%.

Homework: Homework will be assigned approximately every two weeks and will be due on Tuesdays in class. **Late homework will not be accepted.** All homework should be typed (though equations and other mathematical notation can be written by hand). You may work together with other students, but your write-up must be your own. Note: **I will not accept homework sent by email.** Homework will be graded on statistical content and clarity of exposition and presentation.

Policy Projects: The projects should consume the majority of your time. Each group will work with one organization to analyze their data and attempt to answer their research or evaluation questions. You should plan on meeting at least once a week within your project group, and check in at least once a month with your organization. There will also be two required meetings with me (during the weeks of October 9th and November 6th).

During the semester, each group will make brief presentations to the whole class on the progress of their projects. Each group member must participate in each of these presentations. The complete project work will be presented in an end-of-the-year in-class session with the organizations. Here are some important dates:

Milestone	Date	Notes
1st Meeting with Orgs	Week of Sept 25	All group members participate
Practice Presentation 1	Nov 14 & Nov 16	20 min. All group members participate
1st Draft of Paper	Nov 21	Due at 1PM
WA and SPA Weekends	Dec 1-4	Meeting with SPAs and WAs to discuss final papers
Final presentations	Dec 7 & Dec 12	30 min. All group members participate
Final Paper	Dec 12	Last regular course meeting.

Each group must turn in a formal, written report on the last day of class (Tuesday, December 12, 2017). A draft of the written report is due Tuesday, November 21, 2017. Two times during the semester, each student will be asked to assess the contribution of each group member to the team effort including themselves. The organizations will also be given the opportunity to weigh in on each team members contribution. This will be factored into your overall grade.

Responsible Conduct in Research Involving Human Subjects: During the first week of the semester, every student will be required to complete a self-paced, online training program in human subjects research protections. Swarthmore subscribes to the Collaborative Institutional Training Initiative (CITI) which can be found at www.citiprogram.org.

Midterm: Tuesday, October 24 in class
Final Papers Due: Tuesday, December 12th
Final Presentations to Orgs: Thursday, December 7th and Tuesday, December 12 in class

Stat 41: Some Additional Readings

Useful Statistical Modeling Resources:
Seltman, *Experimental Design and Analysis.* **http://www.stat.cmu.edu/~hseltman/309/Book/**
Gelman and Hill, *Data Analysis Using Regression and Multilevel/Hierarchicial Models*, Cambridge
University Press. Some chapters available on Blackboard and also at Haverford
Cook and Weisberg, *An Introduction to Regression Graphics*, Wiley. In Cornell
Weisberg, *Applied Linear Regression* (somewhat more advanced) In Cornell.

References on Statistical Consulting, Scientific Writing, and Professional Ethics:
M. Alley, *The Craft of Scientific Writing,* 3rd edition, New York: Springer-Verlag, 1996. There are
many examples related to the book at **http://www.writing.engr.psu.edu/**
C. Chatfield, "Avoiding statistical pitfalls," Statistical Science 6 (1991): 249-252,
http://www.jstor.org/pss/2245416
D. J. Finney, "Ethical aspects of statistical practice," Biometrics 47 (1991): 331-339,
http://www.jstor.org/pss/2532519
G. D. Gopen and J. A. Swan, "The Science of Scientific Writing," American Scientist 78 (1990): 550-
558, **http://www.americanscientist.org/issues/num2/the-science-of-scientific-writing/1**
W. G. Hunter, "The practice of statistics: The real world is an idea whose time has come," American Statistician 35 (1981): 72-76, **http://www.jstor.org/pss/2683144**
R. E. Kirk, "Statistical consulting in a university: Dealing with people and other challenges,"
American Statistician 45 (1991): 28-34, **http://www.jstor.org/pss/2685235**
R. Tweedie, "Consulting: Real problems, real interactions, real outcomes," Statistical Science 13
(1998): 1-29, **http://www.jstor.org/pss/2676708**
D. A. Zahn and D. J. Isenberg, "Nonstatistical aspects of statistical consulting," American Statistician 37 (1983): 297-302, **http://www.jstor.org/pss/2682767**
J. M. Williams, Style: Toward Clarity and Grace, Chicago: University of Chicago Press, 1990