Syllabus

About the Course

Instructors

- Ben Baumer (`bbaumer@smith.edu`, Burton 315, 413-585-3440). Ben's office hours will be held on Mondays from 2:00-3:30, Thursdays from 10:30-12:00, and by appointment, in Burton 315.
- Amelia McNamara (`amcnamara@smith.edu`, Burton 215, 413-585-3851). Amelia's office hours will be held on Mondays from 10:30-12:00, Thursdays from 10:00-11:30, and by appointment, in Burton 215.

Description

An application-oriented introduction to modern statistical modeling and inference: study design, descriptive statistics, data visualization, random variables, probability and sampling distributions, point and interval estimates, hypothesis tests, resampling procedures, and multiple regression. A wide variety of applications from the natural and social sciences will be used. Classes meet for lecture/discussion with activities and exercies that emphasize analysis of real data.

MTH/SDS 220 is intended to satisfy the basis requirement in introductory statistics for Biological Sciences, Economics, Engineering, Environmental Science, Neuroscience and Psychology majors.

Prerequisite: MTH 111 or MTH 153 or one year of high school calculus, or permission of the instructor. Students who have not taken calculus should consider PSY 201 (Statistical Methods for Undergraduate Research, which also satisfies the basis requirement for the psychology department), SOC 201, GOV 190 or MTH 107 (Statistical Thinking).

Textbook

- *OpenIntro: Introductory Statistics with Randomization and Simulation* (2014), by Diez, Barr and Çetinkaya-Rundel, 1st edition. The textbook is free and open-source; the digital editions can be downloaded for free and paperback copies can be purchased at the bookstore or on Amazon for less than $10. The book is available in three formats:
  - PDF
  - tablet-friendly PDF
  - paperback edition from Amazon

The textbook is not just a reference to use after the instructor has presented new material but a sourcebook to use at every stage of learning. When all students read the text before class, the nature of the class meeting changes to the benefit of everyone. You will have thought about the material, and you will arrive with your own questions. You'll be ready to discuss what you understand, to clarify what you don't understand, and to hear more on the topic. You need to read the book prior to class as well as review the material after we've discussed it in class.

You will be asked to make and respond to reading notes online, as class time will focus on activities to help understand concepts. You should also bring the book to class (in either electronic or paper form), as I will be referencing it regularly.

Classes

Classes meet Monday, Wednesday and Friday in Sabin-Reed 301. Lab meetings are in the same classroom. We expect you to attend class. Your participation is an important part of the learning process. If you cannot attend a particular class we would appreciate the courtesy of advanced notice and an explanation for your absence. Class participation and attendance contribute 5% to your final grade.

We hope it goes without saying, but while the class is in session, you should not use your computer or cell phone for personal email, web browsing, Facebook, or any activity that's not related to the class. Please be sure to bring your book and laptop to class (we'll need to have sufficient numbers of both to allow students to work in pairs).

Policies

Attendance

Your attendance in class is crucial, as is your punctuality. We are all going to learn this material together, so we need to have everyone present and working. We will make accommodations for an unavoidable absence if you notify us. One necessary absence during the semester is not unusual; having more than two is uncommon.

Collaboration

Much of this course will operate on a collaborative basis, and you are expected and encouraged to work together with a partner or in small groups to study, complete homework assignments, and prepare for exams. However, **every word that you write must be your own**. Copying and pasting sentences, paragraphs, or blocks of code from another student is not acceptable and will receive no credit. No interaction with anyone but the instructors is allowed on any exams or quizzes. All students, staff and faculty are bound by the Smith College Honor Code, which Smith has had since 1944.

Academic Honor Code Statement

> Smith College expects all students to be honest and committed to the principles of academic and intellectual integrity in their preparation and submission of course work and examinations. Students and faculty at Smith are part of an academic community defined by its commitment to scholarship, which depends on scrupulous and attentive acknowledgement of all sources of information, and honest and respectful use of college resources.

Cases of dishonesty, plagiarism, etc., will be reported to the Academic Honor Board.

Assignments

1. Homework [15%]: Homework is the most effective way to reinforce concepts learned in class. There will be weekly homework assignments, generally due on Monday. Often, questions will relate to material in the reading that will be covered in class. Homework is due at the start of class, and will be accepted with a 25% penalty if turned in within 48 hours (and no credit otherwise).

2. Labs [10%]: Labs provide the opportunity to delve into real data sets and build your computational and analytical skills. All labs will utilize the statistical programming language R. Lab reports will be written in R Markdown and the resulting HTML files will be submitted via Moodle, generally on Wednesday.

3. Project & Presentation [25%]: Each group of three students will complete a research project during the term, and you will present your results in a final report and oral presentation. Your group will conduct a statistical investigation of a question of interest to you. Rather than collect primary data, you will use data available on the Internet or from faculty research. You will prepare a project proposal describing your study and obtain approval from your instructor before you begin the investigation. During the last week of class, you (and your group) will give a 10 minute oral presentation of your study. The project will give you experience planning a statistical study, acquiring data, creating and testing a linear model, and writing a technical report. We'll talk a lot more about the project as the semester proceeds.

4. Exams [45%]: There will be three self-scheduled, closed-book exams. You will want to bring a scientific calculator and you may bring a double-sided, handwritten piece of paper as a cheat sheet.

5. Participation [5%]: Active participation in class, engagement with the annotated reading notes, and regular attendance will comprise the remainder of your grade.

6. Extra Credit [?]: Extra credit is available in several ways: attending an out-of-class lecture (as will be announced) and writing a short review of it; pointing out a substantial mistake in the book or a homework exercise; drawing our attention to an interesting data set or news article; etc. The extra credit is applied when a student is near the boundary of a letter grade.

Grading

When grading your written work, we are looking for solutions that are technically correct and reasoning that is clearly explained. Numerically correct answers, alone, are not sufficient on homework, tests or quizzes. Neatness and organization are valued, with brief, clear answers that explain your thinking. If we cannot read or follow your work, we cannot give you full credit for it.

Resources

Course Website and Moodle

The course website will be regularly updated with lecture handouts, project information, assignments, and other course resources. Lab assignments and grades will be submitted via Moodle. You should check both regularly.

Computing

The use of the `R` statistical computing environment with the RStudio interface is thoroughly integrated into the course. You have two options for using `RStudio`:

- The **server** version of `RStudio` on the web. The advantage of using the server version is that all of your work will be stored in the cloud, where it is automatically saved and backed up. This means that you can access your work from any computer with a web browser (Firefox is recommended) and an Internet connection.
- A **desktop** version of `RStudio` installed on your machine. The downside to this approach is that your work is only stored locally, and you will have to manage your own installation.

Note that you do not have to choose one or the other – you may use both. However, it is important that you understand the distinction so that you can keep track of your work. Both `R` and `RStudio` are free and open-source, and are installed on most computer labs on campus. Please see the Resources page for help with `R`.

Unless otherwise noted, you should assume that it will be helpful to bring a laptop to class. It is not *required*, but since there are only three workstations in the classroom, we will need a critical mass (i.e. at least 12) computers in the classroom pretty much everyday.

Writing

Your ability to communicate results, which may be technical in nature, to your audience, which is likely to be non-technical, is critical to your success as a data analyst. The assignments in this class will place an emphasis on the clarity of your writing.

Extra Help

There are Statistics TAs available from 7:00-9:00pm on Sunday–Thursday evenings in Burton 301. In addition, the Spinelli Center for Quantitative Learning (2nd Level of Neilson Library) supports students doing quantitative work across the curriculum, and has a Statistics Counselor available for appointments. Your fellow students are also an excellent source for explanations, tips, etc.

Tentative Schedule

The following is a brief outline of the course. Please refer to the complete day-to-day schedule for more detailed information.

Week	Reading	Topic
1	Ch. 1	Introduction to Statistics
2	Ch. 1	Center, Shape, and Spread
3	Ch. 5, 6.1	Linear Regression
4		**Exam 1: 9/29, 1pm until 10/4, library close**
5	Ch. 2	Foundations of Inference
6	Ch. 2	Normal Distribution
7	A	Probability
8		**Exam 2: 10/27, 1pm until 11/1, library close**
9	Ch. 3	Inference for proportions
10	Ch. 4	Inference for a mean
11	Ch. 5-6	Inference for Regression
12		**Exam 3: 12/1, 1pm until 12/6, library close**
13		Projects
12/11		**Project Presentations**
12/14		**Project Presentations**
12/21		**All work due**