Syllabus

About the Course

Instructor

- Randi Garcia (`rgarcia@smith.edu`, Bass 415, 413-585-3698). Randi's office hours will be held on Tuesdays from 1:00p-4:00p, and Fridays from 10:00a-12:00p, or by appointment, in Bass 415.

Description and Goals

This course provides students with an overview of statistical methods needed for scientific research. Our discussions will focus primarily on the basic principles of the design of experiments and observational studies, standard balanced designs, extensions of these designs, and the analysis of data collected under these designs. Most physical, biological, and social processes, produce variable results. This variability can often be quantified and decomposed in ways that enhance our understanding of the process and facilitate decision making. Statisticians and Data Scientists use four steps to quantify and interpret this variability. The steps are:

1. Formulate a statistical question about the process,
2. Design a data collection procedure, or comparison scheme, to answer the question,
3. Collect and analyze the data, and
4. Interpret and communicate the results in written, visual, and oral forms.

One major objective of this course is to give you practice with all four of these steps. Developing our statistical thinking around the factors that produce variability in observations in a key goal of all statistic courses. SDS/MTH 290 course goals:

1. Distinguish between observational and experimental studies, and explain the advantages and disadvantages of each.
2. Explain the three design principles: random assignment, blocking, factorial crossing.
3. Identify the factor structures of the four basic experimental designs: 1) One-way Basic Factorial, 2) One-way Complete Block, 3) Two-way Basic

Factorial, and 4) Split Plot/Repeated Measures.
4. Recognize potential sources of confounding and selection bias.
5. Combine the four basic designs to recommend, recognize, and create new designs.
6. Carry out, document, and explain a randomization plan for a survey or an experiment.
7. Analyze data using R statistical software:
    i. Name and be able to check the six Fisher Assumptions.
    ii. Calculate and interpret an analysis of variance (ANOVA) table for each of the four basic designs.
    iii. Check for and interpret interactions between variables.

Prerequisite: An introductory statistics such as SDS 220/201, ECO 220, PSY 201, or equivalent.

Readings

- *Introduction to Design and Analysis of Experiments*, by George Cobb. Hardcover edition published by John Wiley & Sons,1998, ISBN: 978-0-470-41216-9.

A paperback edition was printed in June 2008. The bookstore has copies of the textbook, and you can also check on-line for used copies or to rent the book electronically. The first few chapters are provided as PDFs on our course Moodle page. It is very important that you read the textbook before every class—time in class will be devoted to jumping right into using the ideas presented in the text. The textbook is written in a conversational style you may enjoy reading! It is full of great examples of real statistical studies.

Class Structure

All class sessions will be focused on active learning activities and there will be less lecturing. We will spend class time discussing your questions, looking at other examples, and doing activities. With this semi-flipped class structure, it is VERY important that you do the assigned readings before coming to class. To this end, I will be starting most class sessions with reading "quizzes" to encourage you to keep up with the reading and to synthesize your thoughts before we begin class (see reading free-write in the assignment section). On some class sessions, I will be giving mini-lectures, but they will often be focused statistical computing in R. Many activities will be designed to give you experience using statistical software to do the extensive computations most statistical analyses require. The computer is faster and more accurate than we are at doing arithmetic and graphs, but we have to know what arithmetic and which graphs will be useful.

The primary way you will get information about this course is from the course website. The course website will be regularly updated with handouts, project information, assignments, and other course resources. All assignments will

be submitted via Moodle. Finally, this course will use Slack to communicate including announcements, discussions about extra credit talks, group project channels, and other uses as the semester progresses.

Policies

Attendance

Participation has two components: 1) being present and 2) engaging in class activities. Your participation is an important part of the process of learning statistics, and we need you in class to help stimulate discussion. You can make a valuable contribution to the discussion by bringing up connections between our work and study designs you have seen in other courses, in the newspaper, or in research literature. I realize that you won't always be able to get to class, and I allow one free absence before I start deducting from your grade (see assignment section below) due to absence.

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

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 instructor is allowed on any exam. All students, staff and faculty are bound by the Smith College Honor Code, which Smith has had since 1944. Cases of dishonesty, plagiarism, or other infractions, will be reported to the Academic Honor Board. You must always provide appropriate citations for others' work and ideas. Giving other scholars due credit in your written and oral communication is a fundamental social gesture in academic work—a way for us to acknowledge each other's scholarship and signify that we respect each other.

From the Smith honor code website:

> 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.

Classroom Environment

Realizing the benefits of a diverse space can only occur if we create a climate of psychological safety (Edmondson, 1999). To this end, we will always be respectful of one another. Together we should have the goal of creating an environment where we all feel comfortable sharing our thoughts and opinions. To this end, I value **"half-formed,"** informal thoughts—sometimes a deeper understanding is reached via communicating ideas before they are perfectly polished.

Please let me know **your desired gender pronoun**. In your written work for this class I am fine with (even encourage) the use of "they," "their," or simply "she" instead of "his or her" or "he or she." I am also fine with "ze" and "zir." Just please do not write "he," "his" or "himself" when referring to all people. We also should also not say "you guys" when referring to a mixed-gender group, or refer to women as "girls."

Accommodations

Everyone should have all that they need to succeed in this course. Please send me your accommodation letter, or have the Disability Office work with me. If you need to register for accommodations, please contact the Disability Services office at ODS@smith.edu. Please check out the office website for more information. Smith provides flexible and individualized accommodation to students with documented disabilities that may affect their ability to fully participate in course activities or to meet course requirements. To receive accommodation services, students must be registered with the Smith College Disability Services office, College Hall 104; phone: (413) 585-2071 (voice, TTY, TDD). Students with disabilities are encouraged to contact the instructor for a confidential discussion of their individual needs for academic accommodation.

If you have special needs concerning test taking, please bring documentation of your needs and make an appointment to discuss them with me, at least TWO WEEKS BEFORE the first test. That will give me time to provide accommodation for your needs.

Assignments

1. **Homework** [20%]: Weekly homework will be due on Wednesdays by 5:00pm in the envelope in my office mailbox (Bass 415). Late homework will NOT be accepted except in cases of a family or personal health emergency. Homework will be a combination of problems sets from the book and statistical computation exercises in R. The computation portion of the homework will often be started in class. I recommend that you form study groups of 3 or 4 students from the class, and get together outside of class to discuss the homework. Each of you should try the exercises on your own and then get together to discuss your work. This process helps you to develop your own way of thinking about statistics questions before hearing how others think. **I will be dropping the lowest homework grade at the end of the semester.**

2. **Reading Free-Writes and Participation** [10%]: Active participation in class and regular attendance will comprise 10% of your grade. Most class periods will begin with 5-minute 'open-book' reading free-writes. By open-book I mean that you may flip through the book to remind yourself—there will be no time to actually read during the free-write. This process will help us move towards a "flipped" classroom. In non-flipped classrooms, students attend lectures and are asked to apply the material outside of class in the form of homework assignments. But often students need the most help from professors when they are in the midst of applying the material (Bergman & Sams, 2012). In flipped classrooms, students are asked to learn the facts at home and do the application in class.

   You will need to introduce yourself to the chapter material independently and complete the reading free-writes before the material is formally presented in class. The reading free-writes will often ask you to highlight the most important chapter material and will aid you in determining what concepts you might need to ask questions about in the class period. We often don't know what we don't know, until we try to write about it! Content will be flexible—sometimes I will ask you to describe specific statistical concepts, on other days I may ask you to simply write 3 things that you are still fuzzy on with regards to the reading. Grading for these reading free-writes will be on a 3-point scale (with the possibility for an extra credit point): 0 – not completed, 1 – completed but clear that you didn't read, 2 – completed and clear that you read, and (3 – extra credit for an especially insightful free-write). You can miss up to two free-writes without penalty to your grade.

3. **Exams** [40%]: There will be two take home exams usually distributed on the Wednesday and due that next Monday. The exam format will be in two parts: 1) written answers to questions (handwritten or in R Markdown) and 2) the analysis of a dataset (in R Markdown). Both parts are due at class time on the Monday due date.

4. **Group Project and Poster Presentation** [30%]: Each group of four to five students will complete a research project during the term. We will be collaborating with the Smith Botanic Garden on these projects. Your group will be prosing your own research questions (in consultation with the Botanic Garden staff) and designing an experiment (submitted as a short proposal due Feb 19). Tim Johnson, Director of the Botanic Garden, will visit our class on Feb 7 and we will have a class field trip to the Botanic Garden on February 12th. The specific content of your projects will vary, but all projects will consist of designing an experiment (or appropriate observational study), collecting and analyzing data, and writing a technical report on your study. During the last two weeks of class, you (and your group) will create a research poster and participant in a poster presentation of your study. We'll talk more about the project as the semester proceeds and detailed instructions for the project will be posted on the course website.

5. **Extra Credit** [?]: Extra credit is primarily available by attending an out-of-class lectures (acceptable lectures will be announced) and writing a short review of it on Slack; participating in a research study; etc. The extra credit is applied when a student is near the boundary of a letter grade.

Summary

Assignment	Percent	DueDate
Homework	20%	Wednesdays
Reading Quizzes	5%	Daily
Participation	5%	Daily
Take Home Exam 1	20%	March 19th
Take Home Exam 2	20%	April 23rd
Group Project		
Group Project Proposal	5%	February 19th
Group Project Update/Method Draft	5%	March 5
Final Poster and Presentation	10%	May 2
Final Technical Report	10%	May 9
Final Grade	100%	

Grading

When I grade your written work, I am looking for problem solutions that are technically correct and reasoning that is clearly explained. Numerically correct answers, alone, are not sufficient on homework or exams. I value neatness and brief, clear, well-organized answers that explain your thinking. If the grader cannot read or follow your work, she cannot give you credit for it. The grader

will check each homework submission for completeness and grade a subset of the exercises. Homework answer keys will be posted on Moodle after the due date.

Assignment Late Policy

For every day late on the written assignments I will reduce your grade for that assignment by 1/3 letter grade (i.e., if you wrote a B+ paper, but turned it in 1 day late, you'll get a B. Two days late? B-, and so forth). I will not accept any late final group research papers.

Final Grade Brackets

Grade	Percent
A	95-100%
A-	90-95%
B+	87-89%
B	83-86%
B-	80-82%
C+	77-79%
C	73-76%
C-	70-72%
D+	60-66%
D	67-69%
E	59% and below

Resources

Course Website and Moodle

The course website will be regularly updated with lecture handouts, project information, assignments, and other course resources. 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.

Use of Technology in Class

We will be using laptops (and sometimes smart phones) extensively in this course. Please bring them with you to every class period. If you do not have a laptop, you can loan one for the semester via ETS's laptop loan program. Alternatively, there are a few desktop computers available in the classroom that can be used if need be. I hope it goes without saying that while the class is in session, you should resist the temptation to use your computer or cell phone for personal email, web browsing, social media, or any activity that's not related to the class.

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 McConnell 301. In addition, the Spinelli Center for Quantitative Learning (Seeley 207) 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 Four Basic Designs
2	Ch. 2, 4	Decisions about Content of an Experiment
3	Ch. 3	Formal Analysis of Variance (ANOVA)
4	Ch. 5	Basic Factorial Design
5		**Exam 1: take home due on March 19th**
6	Ch. 6	Factorial Crossing
7	Ch. 7	Complete Block Design and the Split-Plot Design
8	Ch. 9	Extending the Basic Designs
9	Ch. 11, 12	Correcting for Multiple Comparisons and Assumption
10	Ch. 13	Nesting; Fixed and Random Effects; Expected Mean S
11		**Exam 2: take home due on April 23rd**
12		Projects
13		Projects
5/2		**Poster Presentations**
5/9		**All work due**