

## STAT 113: INTRODUCTION TO STATISTICS (FALL 2017)

**Instructor.** Colin Reimer Dawson (*he/him/his*)

**Office.** King 204

**Email.** cdawson@oberlin.edu

**Website.** <http://colindawson.net/stat113/>

**Office Hours.** M 1:30-2:30, T 4-4:50, Th, 4-5:30, F 9-9:50 and 11-11:50 *If your schedule prevents you from attending any of the posted hours, please let me know as soon as possible.*

**Times and Locations.** T/Th classes are 9:30-10:50 in King 241 and F lab is 10-10:50 in King 137.

### COURSE OVERVIEW

**General Learning Outcomes.** By the end of this course you should be equipped to use data to evaluate (with the help of a computer) the strength of evidence for simple claims (e.g., “homeopathic headache medicine reduces pain more effectively than placebo”; “parental income is a strong predictor of SAT scores”).

Note that this is a *statistics* course and *not* a math course. The point of the course is *not* learning how to calculate things, nor is it memorizing terms and definitions. It is not even primarily about understanding the mathematics that supports the statistics. It is about understanding what data is (and isn’t!) telling you, understanding *why* it is telling you that, and communicating about that process.

Some of the concepts in statistics can be quite subtle, even while the calculations involved are relatively straightforward. We will devote most of our energy (in class, on homework, on exams) to understanding those nuances, and examining the thought processes involved.

**Who Should Take This Course.** This is intended as a first course in statistics, and does not assume any previous exposure to the subject. It is likely to be a useful course even if you have taken statistics in high school, since we will be delving deeper into the concepts than most high school AP classes; however, if you have a score of 3 or above on the AP stats exam, you may want to take STAT 215 instead, which starts in the same place but moves at a faster pace.

---

*Date:* Last Revised September 9, 2017.

**Grading and Specific Learning Objectives.** In my ideal world, there would be no grades and we would just work on learning cool stuff for the intrinsic benefit of learning. But grades can be useful as a bit of extrinsic motivation, and also as something to show others that you did something. However, I hope to minimize the focus on grades as evaluation, and maximize the focus on learning with grades as a feedback mechanism. With that goal, the grading system in this class is non-traditional, and takes some explanation. See the handout called “Grading System” for a description of specific learning objectives (SLOs), and for details about the grading system and of grade-related course policies.

**Lab.** Throughout the semester, the lab component of the class will introduce computational techniques (using the statistical language R) to offload the tedious work of counting, arithmetic, plotting, etc. However, while this offloading frees us to focus on conceptual understanding, it is important to note that being able to translate your goals into explicit terms that a computer can understand is an important skill in itself, to which you will likely devote a significant chunk of time and effort. Although the specific commands you learn in this class are just one way of doing what they do, and will likely be superseded in the near future as new and better software is developed, my hope is that you will learn to *think computationally*, which is an incredibly useful skill in this day and age that goes beyond this class.

## MATERIALS

**Textbook.** The textbook is *Statistics: Unlocking the Power of Data*, by Lock, Lock, Lock, Lock and Lock (it’s a “family textbook”), or “Lock 5” for short. Section numbers and exercises refer to the second edition, which is an improvement, in my opinion, over the first edition. There are both physical and electronic versions available from the Oberlin book store. We will cover almost the whole book. If you intend to buy a physical copy but are waiting for it to arrive, you can get a free 7-day trial of the electronic version of the first edition on Amazon.

**Software.** We will use the free statistical computing environment RStudio, which is an interface to the language R. You may either install R and RStudio on your personal machine ([www.r-project.org](http://www.r-project.org) and [www.rstudio.com](http://www.rstudio.com), respectively), or, once your account is set up, use Oberlin’s RStudio server via a web browser ([rstudio.oberlin.edu](http://rstudio.oberlin.edu)). We will talk about pros and cons of each in lab. The R language has become the standard computing tool used by practicing statisticians and data scientists, and so although statistical reasoning is the main goal of the course, competence in R and written presentation of results is a learning objective unto itself as well.

## MISCELLANY

**Laptops.** We will do some group activities involving the computer during some non-lab classes, so if you have a laptop you may want to bring it to class (only one person per group will need one). Some activities can be done on a tablet or smartphone instead. However, I strongly discourage the use of electronic devices outside these activities. For note-taking, take hand-written notes if you are physically able to do so: there is evidence that writing by hand improves your cognitive processing of the material.

**Email.** Email is the best way to reach me outside of a face-to-face meeting. You are welcome to address me by my first name, which is generally what I will use when signing emails.

I try to respond to most email within 24 hours. If I have a lot, and it is not pressing, as well on the weekend, it might be longer. **If you need to ask me about something due the following morning, don't wait until the night before.** I have limited work time in the evenings due to family obligations, and besides, it's just poor form.

**Accommodations.** If you have a disability of any sort that may require accommodations in order for you to do your best work in this class, please let me know as early as possible, and consult as well with the Office of Disability Services (ODS). By college policy, **all requests for accommodation require documentation from ODS.**

**Honor Code.** The Oberlin College Honor Code formalizes the idea that all work that you submit is your own and that you have given credit to the ideas and work of others when you incorporate them. You will be asked to write and sign the honor pledge on each graded assignment that you hand in. The honor pledge reads: "I have adhered to the Honor Code in this assignment."

What it means to adhere to the honor code depends on context. For each assignment type, I describe what it means to follow the honor code on that assignment below.

More information about the honor code can be found on the web at the Dean of Students site:

<http://new.oberlin.edu/office/dean-of-students/honor/students.dot>

## YOUR RESPONSIBILITIES

**Readings.** There will be daily reading assignments posted on the website. *It is important to do the readings*, either before or soon after material is discussed, especially since some topics will not be covered in a “lecture”. Occasionally you may not have done the reading. When this happens, just be up front about it. As long as it’s not too often, no one will hold it against you.

**Attendance and Participation.** I will spend some class time presenting material, but we will devote much of our time to explorations of the material in groups. There is consistent evidence that learning via self-discovery is more effective than passively listening to a lecture. There is no formal attendance grade, but I expect you to attend class nearly always and participate actively. I will take attendance and participation into account when making judgment calls about borderline final grades.

**Homework and Labs.** There will be (roughly) weekly assignments, posted on the course website (not Blackboard), mostly consisting of textbook problems. These will be due electronically via Blackboard, generally by class time on Tuesdays (see the course website for a specific schedule of due dates). A few of these problems may count toward the grade; others will be looked at and given feedback but will not be used in grade calculations. There will also be lab assignments that will be worked on during the Friday lab, with additional homework questions to finish on your own. In most cases, these will be due together with the week’s homework assignment on the Tuesday after the lab. Details on the homework policy are in the “Grading Policy” handout (which is also on the website).

*Honor Code:* I encourage you to work on homework and labs collaboratively; however, **you must write your own solutions and code; you may not copy each other’s words or commands.**

**Mini-Project.** There will be one guided “mini-project” that will take the place of one of the labs toward the end of the semester. It is mostly like a lab, but it is more open-ended, more integrative, and assesses more learning objectives.

*Honor Code:* Same as for labs

**Quizzes and Exams.** There will be a short (~10 min) quiz at the start of class most Thursdays (excluding the first week and weeks when there is an exam). There will be three in-class exams evenly spaced throughout the semester. There is no final exam during finals period; however, we will convene during the scheduled final exam time for presentations of final projects. Exams may only be made up in the event of

a serious emergency. Quizzes may not be made up under any circumstances, but see the policy on reassessments in the “Grade Policy” handout.

*Honor Code:* Exams and quizzes must be done individually. The exams will be closed book, but calculators and a single-sided, *hand-written* 8 1/2” × 11” note sheet will be allowed. The requirement that your sheet be handwritten is there to encourage you to condense and process material while creating your sheet, so that creating the note sheet is itself a study opportunity.

**Term Project.** There will be a data analysis project with milestones throughout the semester, to be done in teams of 2-3, culminating in a 5-7 minute group presentation and a final paper at the end of the semester. The presentations will occur during the appointed final exam time; the final paper is due at the same time (though the option to do minor revisions in light of feedback received at the presentation may be available). The purpose of these projects is to identify a question about a topic of interest to you, find or collect data that can shed light on your question, and use the concepts and techniques you learn in class to try to answer the question.

*Honor Code:* Free collaboration with your team members is of course required. Describing the work that other people did before you is an important part of intellectual inquiry, and you must give credit and cite sources for for any data or ideas that did not originate within your team. This includes paraphrases as well as direct quotations. All members of the group must make approximately equal contributions at each stage of the project.

#### KEY DATES

**Final Presentations.** Friday, 12/15, 9-11 A.M., in lieu of a final exam. Arrive by 8:45 for bagels and coffee on me.

**Hour Exams.** Three in-class exams are tentatively on 10/6, 11/3 and 12/8. These are Fridays and will be in the computer lab. The dates may change based on adjustments to the pace of the class, or unusually many conflicts (e.g., for athletic travel). **Let me know by mid-September if you have a pre-arranged conflict with any of these dates** so that if needed we can rearrange well in advance.

**Holidays.** The only college holidays that affect our T/Th/F schedule are fall break (10/16-10/20) and Thanksgiving (11/23-11/24).