MATH/LING 289: Introduction to Data Science, Fall 2018
Tuesday, Thursday 12:00-13:15   MRC LL1

Instructor: **Taylor Arnold**
E-mail: tarnold2@richmond.edu
Office: Jepson Hall, Rm 218
Course Website: `https://statsmaths.github.io/stat289-f18`

Overview:

Data science is an interdisciplinary field concerned with drawing knowledge from data and communicating those results to various audiences. Data science needs to be learned *by doing* data science. There are no short-cuts and the process cannot be learned by simply working our way methodically through a textbook of disconnected topics.

Therefore, this course will be taught using a problem-based learning model. As a class, we will be addressing an open ended research question and discussing various methods that will assist our inquiry. Both individually and in small groups, students will take responsibility for particular subtasks that drive our research forward. At the end of the semester students will have acquired a toolkit of methods, and the knowledge of how to use them in practice, to address important social, cultural, and scientific questions with data-driven techniques.

We will make heavy use of computing throughout the semester, but *no prior programming experience is required*. Also note that this course has a MATH designation because statistics is currently housed in the mathematics department. The topics of this course, however, do not fall within the traditional disciplinary boundaries of mathematics.

Course Website:

All of the materials and assignments for the course will be posted on the class website:

> `https://statsmaths.github.io/stat289-f18`

The website contains notes, assignment details, and supplemental materials. At the end of the semester, this version of the course will be archived and available for your reference.

GitHub:

All of your work for this semester will be submitted through GitHub, the same platform that hosts our website. You'll need to set up a free account, which we will cover during the week of class.

Labs:

During most class meetings in the first half of the semester will be spent working together through IPython notebooks. I'll call these 'labs' for the purpose of this syllabus. Each lab consists of a set of questions that must be answered with either small snippets of code or short descriptive answers. We will largely work on these together, but often the final bit will need to be finished on your own. Solutions must be uploaded to your GitHub page prior to the start of the next class meeting. Labs are graded on a Pass/Fail scale.

Exams:

There will be three take-home midterm exams given throughout the semester. These are open notes and resources, though you must complete them on your own. The scheduled dates for the exams are (all dates are Thursdays):

- **2018-09-20**
- **2018-10-18**
- **2018-11-15**

The exam will be made available starting at the normal start of class on the date above; you will have 5 days (starting at the normal class start time on Thursday and continuing through start of class on the following Tuesday) to complete it. The course has no final exam.

Final Project:

The final project for this class will consist of either a small software library or a written data analysis report. More details will follow about the specific requirements and deadlines.

Final Grades:

The final grade will be determined by weighting the labs and projects as follows:

- **Labs and Participation**: 5%
- **Midterm Exams**: 45% (15% each)
- **Final Project**: 50%

To pass the course, you must also miss no more than four class meetings. Attendance requires that you arrive on-time, complete any out of class assignments for the day, and fully engage with the course material. Failing to fulfil these attendance requirements may result in a failing grade of 'V' or a reduction to your final course average at the instructor's sole discretion.

Class Policies:

The following class policies address some of the most common questions and concerns that students have. If anything is unclear, please feel free to contact me for clarification at any point in the semester.

- **Academic honesty:** Cheating and plagiarism are grave scholarly offenses and potential grounds for expulsion; they are also a major barrier to your intellectual development. You are expected to familiarize yourself with the entirety of the University of Richmond's Honor Code. If you are confused or unsure about appropriate citation protocol or any other aspect of the Honor code, please consult me before turning in an assignment.

- **Special approval:** If you have special approval forms for extra time on exams or any other circumstances I should know about, please speak with me as early as possible so that we can best accommodate your needs.

- **Late work:** You are expected to submit all work on-time. Late work will be accepted after the due date with a full letter grade deduction for each 24 hour period it is late (rounded up).

- **Attendance:** You are expected to both attend and participate in most class meetings. If you must be absent due to illness or other pressing need, please let me know by email as soon as possible. A habit of arriving late, failing to participate, or failing to accomplish any out of class assignments is considered equivalent to an absence.

- **Make-up work:** In instances where students have a valid excuse for missing an assessment, please get in touch with me within 24-hours of missing class to make alternative arrangements.

- **Class conduct:** During class I expect you to refrain from checking email, being on phones, or working on assignments for other classes.

- **Computers:** I except you to bring a working laptop with Python 3 installed. If this poses a challenge, please speak with me at the start of the semester or as new problems arise.

- **Office hours**: Rather than fixed weekly office hours, I will provide blocks of open times to meet with me particularly focused around exams and project due dates. If you find me in my office, poke your head in and I am usually happy to meet on the spot. Otherwise, please email me to make an appointment so that we can chat. Please note that appointments should be booked at least 24 hours ahead of time.

- **Email:** I will also answer questions by email (it can, in fact, be much faster than scheduling an appointment for small issues). During the week, I aim to respond within 24 hours, with emails sent over the weekend responded to by Monday morning. If your question involves code, please attach your current lab or report as that will expedite my answering your question(s).

Notice:

I reserve the right to modify this syllabus, with advanced warning, throughout the semester. If necessary, I will email the class list and post an updated version of the document on the course website.