MATH/STAT 209: Introduction to Statistical Modeling, Fall 2018
Tuesday, Thursday 10:30-11:45   MRC LL1

Instructor:	**Taylor Arnold**
E-mail:	tarnold2@richmond.edu
Office:	Jepson Hall, Room 218

Description:

This course broadly covers the entire process of collecting, cleaning, visualizing, modeling, and presenting datasets. The focus is on applied statistics and data analysis rather than a detailed study of symbolic mathematics. It has a MATH designation but is not a *mathematics* course. By the end of the semester you will feel confident collecting, analyzing, and writing about datasets from a variety of fields. You will be able to use these skills to address data-driven problems in a wide range of application domains.

Course Website:

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

> `https://statsmaths.github.io/stat209-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.

Computing:

To facilitate your ability to actually *do* statistics, most class meetings will involve some form of computing. No prior programming experience is assumed or required.

We will use the **R** programming environment throughout the semester. It is freely available for all major operating systems and is pre-installed on many campus computers. You can download it and all supporting files for your own machine via these links:

> `https://cran.r-project.org/`
> `https://www.rstudio.com/`

The lab computers in Jepson are available and contain all of the required software. I strongly recommend, however, downloading these on your own machine so that you will be able to work on assignments without needing to work only in the computer lab.

Labs:

Nearly every course meeting will have an associated lab to complete. We will usually work on these in class. By the start of the next course meeting, you must complete the questions contained within the lab notebook. Assignments will be submitted through GitHub; this process will explained in more detail during class.

Midterm Exams:

The course will have three in-class exams throughout the semester. These will resemble the daily lab assignments and will be administered on the computers in our classroom. Dates for these exams will be distributed roughly evenly throughout the first twelve weeks of the course:

- **Midterm I**: 2018-09-20 (Thursday)
- **Midterm II**: 2018-10-23 (Tuesday)
- **Midterm III**: 2018-11-15 (Thursday)

The two weeks of class following Thanksgiving break will be focused on finishing and presenting your final data project.

Data Projects:

While the midterms serve to make sure you are following along with the general concepts, the ultimate aim of the course is to teach you how to *apply* statistics to real-world questions. To this end, you will also complete a final data-oriented project for the course. Details for this project and how it will be graded will be given towards the middle of the term. The final week of the course will consist of presentations of your findings.

Final Grades:

The final grade will be determined by weighting the labs, exams, and project as follows:

- **Labs and Participation**: 20%
- **Midterm Exams**: 60% (20% each)
- **Final Project**: 20%

Course expectations and community standards will be discussed, developed and distributed in the first week of the course. This will include policies for class participation, attendance, and late work.

Weekly Topics:

These are subject to change, but will give you a sense of what will be covered throughout the semester:

- Week 01: Introduction to Data Analysis
- Week 02: Introduction to R and Data Construction
- Week 03: Reproducible Research and Variable Types
- Week 04: Review and MIDTERM I
- Week 05: A Grammar for Data Visualization
- Week 06: Annotating Graphics for Publication
- Week 07: Filtering and Mutating Data
- Week 08: Summarizing Numeric Data
- Week 09: Review and MIDTERM II

- Week 10: Statistical Inference
- Week 11: Joining and Reshaping Data
- Week 12: Review and MIDTERM III
- Week 13: Final Project Consultations
- Week 14: Final Project Presentations

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.