

**MATH/CS 395: Statistical Learning, Fall 2017**  
Tuesday, Thursday 13:30-14:45    PURH G13

**Instructor:**    **Taylor Arnold**  
**E-mail:**        tarnold2@richmond.edu  
**Office:**         Jepson Hall, Rm 218  
**Office hours:**   Tuesday, Thursday 10:30-12:00 or by appointment

**Computing:**

The focus of this course will be on applied statistics and data analysis over symbolic mathematics. To facilitate this, nearly every class assignment and exam 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/>

I strongly recommend using your own machine for this course, though the lab computers in Jepson are available and contain much of the required software as well.

**Course Website:**

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

<https://github.com/statsmaths/stat289>

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.

**Grades:**

All grades in this course will be given on as a letter grade. While occasionally possible to receive pluses / minuses or fractional points, these will usually be given a whole letter grade.

I expect most students to get full marks (A) for labs and participation. Students found to be delinquent in either will first receive a written warning, followed by an initial 50% reduction (C) in the respective grade, and finally a 100% reduction (F).

Your final grade will be determined by converting all grades into a numeric scale as follows (pluses increase the number by 0.33 and minuses decrease the number by 0.33):

Numeric Score	Final Grade
4	A
3	B
2	C
1	D
0	F

I want to make the grading extremely transparent, so your final grade will simply consist of taking your weighted numerical average using the following weights:

- Participation, 20%
- Reports, 80%

And reading off of the following chart (grades are rounded to the second digit):

Numeric Score	Final Grade
3.84 - 4.00	A
3.50 - 3.83	A-
3.17 - 3.49	B+
2.84 - 3.16	B
2.50 - 2.83	B-
2.17 - 2.49	C+
1.84 - 2.16	C
1.50 - 1.83	C-
0.00 - 1.49	F

### **Reports:**

The reports required for this class will be written documents that mix code, graphics, and prose to provide a comprehensive analysis of a predictive modelling task. Reports will be due at the start of class on most Tuesdays. These must be submitted on GitHub. The format for these reports will be described in the first week of the course.

### **Exams:**

This course has no exams, final or otherwise.

### **Participation, Attendance and Late Policy:**

You are expected to submit work on-time. You should aim to attend all class meetings, however I am fully aware that through the course of the semester various issues – illness, sports, and family emergencies – will prevent many of you from attending every class. As long as you come prepared to most meetings of the course you will receive full credit for participation. Likewise, attending every class but continually failing to engage in the classroom activities will not earn full course credit.

**Weekly Topics:**

These topics are subject to change based on the pace of the course, but give a good sense of roughly what we are going to cover:

WEEK 01 - Introduction to R, RMarkdown, and Graphics

WEEK 02 -

WEEK 03 -

WEEK 04 -

WEEK 05 -

WEEK 06 -

WEEK 07 -

WEEK 08 -

WEEK 09 -

WEEK 10 -

WEEK 11 -

WEEK 12 -

WEEK 13 -

WEEK 14 -