STT 3530 (Data Science I) Syllabus

Fall 2017

Instructor: Dr. Alan T. Arnholt
Office: Walker Hall 237
Office Hours: 10-11:30 M, W, and F

Make an appointment to see me by clicking https://arnholtat.youcanbook.me/.

Course Description:

Data science is an exciting discipline that allows you to turn raw data into understanding, insight, and knowledge. The goal of Data Science I is to help you learn the most important tools in R that will allow you to do data science.

Course Objectives:

1. Students will learn how to use a reproducible research work flow.
2. Students will improve their technology expertise.
3. Students will learn to work with large data sets.
4. Students will learn the grammar of `ggplot2`.
5. Students will learn to display their work on the web.
6. Students will learn how to read data into R.
7. Students will learn to store data in a `tidy` format.

Course Texts:

R for Data Science

Course Grading & Assessment:

The only way to learn statistics is to **DO** statistics, which includes statistical software. Reading the textbook, learning the language, and practicing exercises using real data are critical to your learning and success. Class activities and assessments have been structured with these principles in mind.

You should read assigned textbook content and read/watch supplemental materials prior to coming to class. It will be easier to participate if you acquire some familiarity with the vocabulary and methods before we start to discuss and use them. You must "speak the language" (both statistics and R) to effectively demonstrate your knowledge.

Appalachian students are expected to make intensive engagement with courses their first priority. Practically speaking, students should spend about 1-3 hours on coursework outside of class for every hour they spend in class.

- 60% of the course grade will come from the thirteen DataCamp assignments
- 15% of the course grade will be for the design and creation of your web page
- 25% of the course grade will come from the final project

Grades will be kept on AsULearn. You should monitor your grades throughout the term to make sure they appear to be correct and complete. Feel free to inquire about your grades at any time.

How To Get Unstuck

Well constructed questions will elicit answers more rapidly than poorly constructed questions. This video provides some background on asking questions. This stackoverflow thread details how to create a minimal R reproducible example. Please read How To Ask Questions The Smart Way by Eric Raymond and Rick Moen and heed their advice.

University Policies

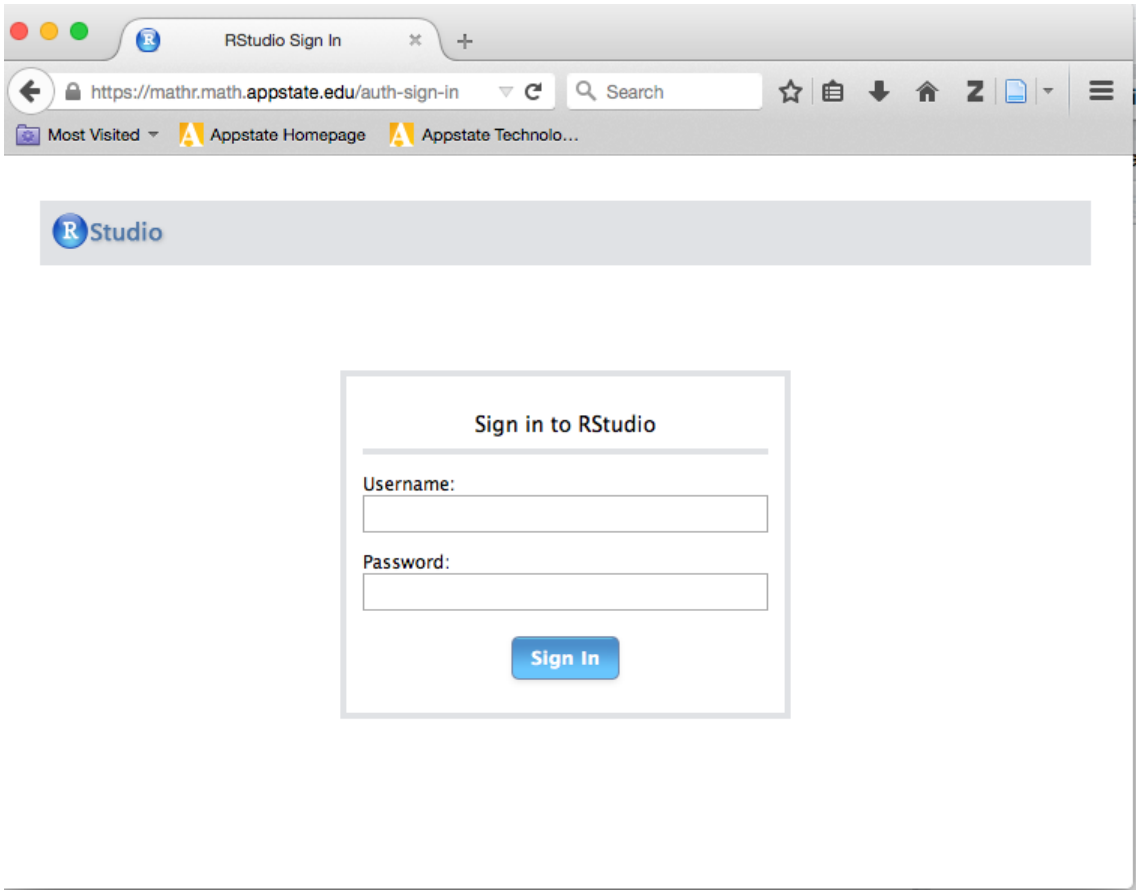
This course conforms with all Appalachian State University policies with respect to academic integrity, disability services, and class attendance. The details of the policies may be found at http://academicaffairs.appstate.edu/resources/syllabi.

Computers and Software

This course will use the RStudio server (https://mathr.math.appstate.edu/) that has the programs listed below and more installed.

- R
- Git
- RStudio
- LaTeX

You must have an active internet connection and be registered in the course to access the server. To access the server, point any web browser to https://mathr.math.appstate.edu/. You will need to acknowledge the connection is unsecure and possibly add a security exception to your web browser. Use your Appstate Username and Password to access the server. A screen shot of the RStudio server is shown below.



If you have problems with your Appstate Username or Password visit IT Support Services or call 262-6266.

Required Technology

- RStudio Server
- DataCamp
- GitHub
- AsULearn

Note: All technology used in the class is either open source (free) or will be accessible to students enrolled in the course for no cost.

Assignments

The CoursePacing guide has all course assignments and due dates.