STT 3851 Syllabus - Fall 2021

Fall 2021

Instructor: Dr. Alan T. Arnholt
Office: Walker Hall 237
Student Help Hours: 1-3pm M & W, 10am-12pm R, and by appointment

Make an appointment to see me by clicking here.

Course Description:

This course covers elements of data management, descriptive statistics, and inferential statistics. The course also examines the variance-bias tradeoff, linear regression, cross-validation, bootstrapping, subset selection, ridge and lasso regression, and choosing optimal models.

Course Objectives:

1. Students will use a reproducible research work flow.
2. Students will improve their technology expertise.
3. Students will learn to apply statistical learning techniques.

Course Texts:

The principal documents for this course are *ModernDive: An Introduction to Statistical and Data Sciences via R* (MD), *Data Science with R* (DSWR), and *An Introduction to Statistical Learning with Applications in R* (ISLR)

Optional References:

Reproducible Research with R and RStudio, Second Edition by Christopher Gandrud

The Elements of Statistical Learning by Trevor Hastie, Robert Tibshirani, and Jerome Friedman

R Graphics Cookbook by Winston Chang - Available via `SafariBooksOnline` through the Appalachian State University library.

Course Grading:

The only way to learn statistics is to **DO** statistics, which includes using 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 demonstrate your knowledge effectively.

Appalachian students are expected to make intensive engagement with courses their first priority. Practically speaking, students should spend approximately 2-3 hours on coursework outside of class for every hour they spend in class. For this three-hour course, you you should anticipate **6-9** hours per week of outside work.

Grade Distribution:

- 25% of the course grade will come from (25) Data Camp assignments
- 40% of the grade will come from four labs and three problem sets
- 20% of the grade will come from reproducing two data camp assignments with bookdown
- 10% of the grade will come from a kaggle competition
- 5% of the course grade will come from a presentation during the final exam period

---

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 face coverings, academic integrity, disability services, class attendance, and student engagement. The details of the policies may be found at https://academicaffairs.appstate.edu/resources/syllabi-policy-and-statement-information. Please pay particular attention to the student engagement statement.


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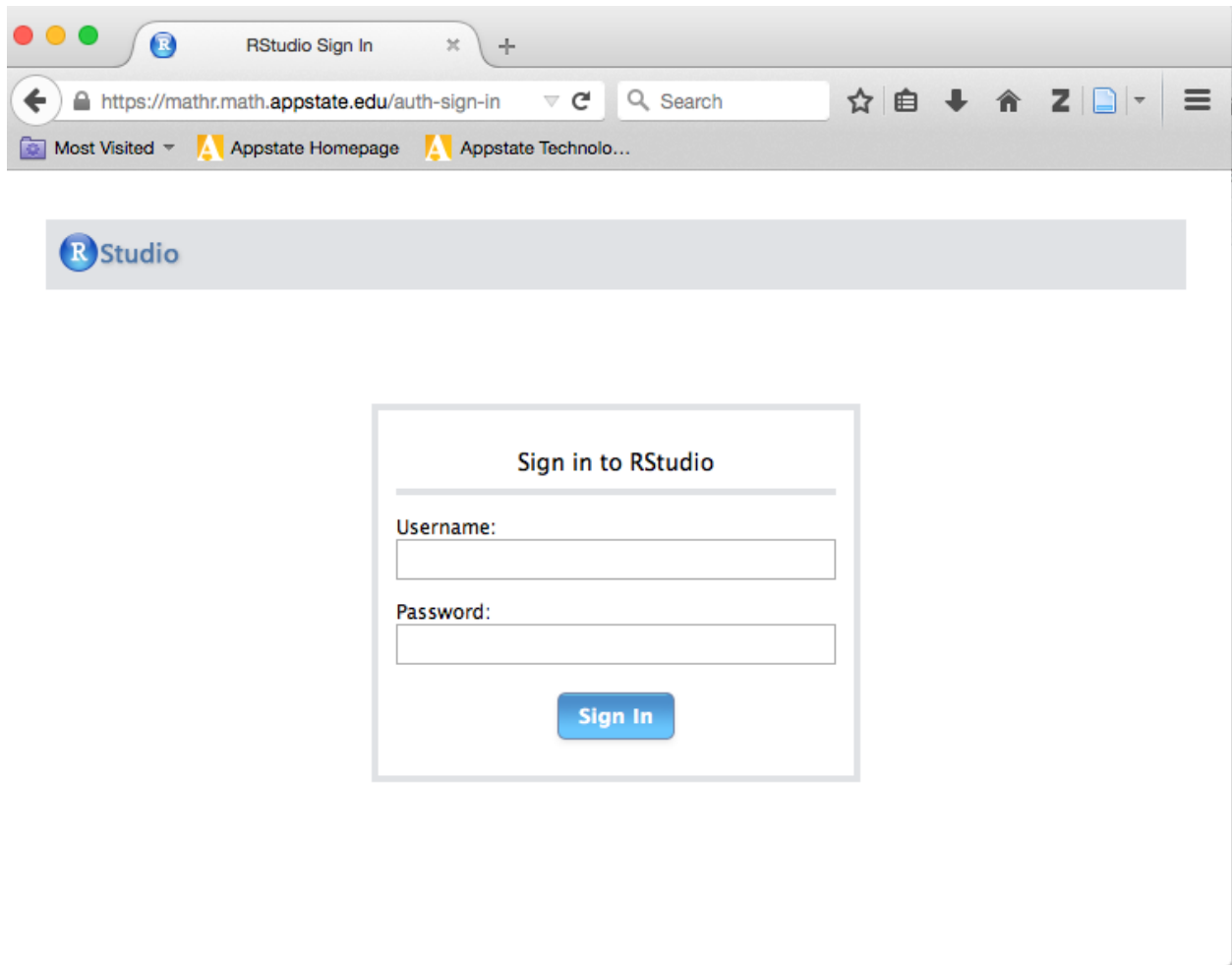
If you have problems with your Appstate Username or Password visit IT Support Services or call 262-6266.