

STT 2860 Syllabus - Fall 2024

Instructor: Dr. Alan T. Arnholt

Office: Walker Hall 237

Student Help Hours: 12:30-2:00 pm T & R, 2:00-3:30 pm W, and by appointment.

Make an appointment to see me by clicking [here](#).

Course Description:

This course is an introduction to the basics of data management and visualization, including reproducible workflow. Topics include introductory concepts of programming and work flow, data manipulation, and data visualization using grammar of graphics. The course will also emphasize the practice of creating reproducible research using a version control system with dynamic document reporting, including tools such as R, RStudio/Posit IDE, Quarto, and Git.

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 Text:

- *R for Data Science* (R4DS)

Supplemental Materials:

- Statistical Inference via Data Science (MD)
- `ggplot2`: Elegant Graphics for Data Analysis (3e)
- R Graphics Cookbook, 2nd edition
- Happy Git and GitHub for the useR

Course Grading & Assessment:

- 30% of the course grade will come from DataCamp assignments (28). **DataCamp** assignments are assigned for students to practice coding and to receive immediate computerized feedback. You should attempt to answer the DataCamp questions correctly and not simply ask the program to show you a solution. 60% of your DataCamp grade will be a binary grade (1 or 0) for completion of the assigned DataCamp Chapter. The remaining 40% of your DataCamp grade will be computed at the end of the course and will be based on your total XP points accumulated during the semester in DataCamp. (XP greater than 25,500 A; XP greater than 24,000 B; XP greater than 22,500 C; XP greater than 21,000 D)
- 20% of the course grade will come from your answers to the assigned Exercises from R4DS.
- 10% of your grade will come from your Web page.
- 20% of your grade will come from the reproduction of the Bike Sharing example.

- 20% of your grade will come from the analysis of an approved data set.
-

University Policies

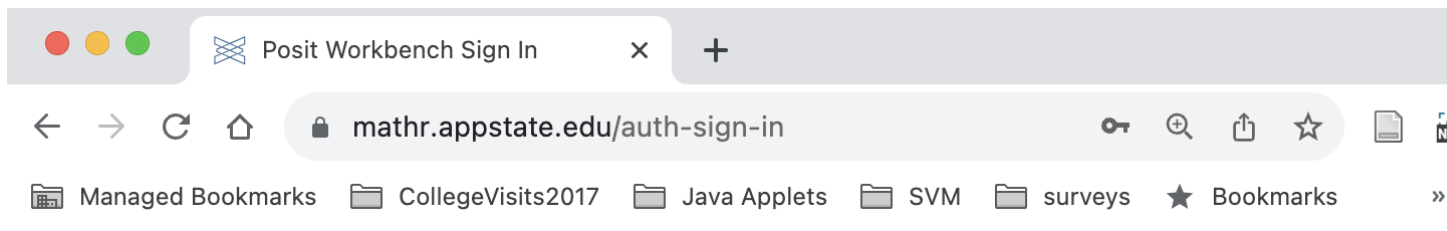
This course conforms with all Appalachian State University policies with respect to academic integrity, disability services, class attendance, and student engagement. The details of the policies may be found at <https://academicaffairs.appstate.edu/resources/syllabi-policy-and-statement-information>. Please pay particular attention to the student engagement statement.

Computers and Software

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

- R
- Git
- RStudio IDE
- 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.appstate.edu/>. Use your Appstate Username and Password to access the server. A screen shot of the POSIT workbench login screen is shown below.



Sign in to Posit Workbench

Username:

Password:

☐ Stay signed in when browser closes

You will automatically be signed out after 60 minutes of inactivity.

Sign in

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

Required Technology

- RStudio Server
- DataCamp
- GitHub
- Github Classroom Repository

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 Course Pacing Guide has all course assignments and due dates.

Faculty student responsibilities

- It is my (faculty) responsibility to explain and present the material you need to master for this course. A detailed description of everything you need to do starting with day one to the Final Exam is provided in the course pacing guide which is available on day one of the course.
- It is your (student) responsibility to learn the material and to seek help if you do not understand the material.

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 should anticipate 6-9 hours per week of outside work.

How To Do Well In This Course

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. When you read the assigned material, you should complete the problems (not just read about them) on your paper and computer. 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. If you come to class and have difficulty following the discussion, you should make sure you have read all of the assigned material and then go back and re-read the assigned material a second time. Reading a technical book is not the same as a novel. Most people, your instructor included, must read a technical section at least twice before understanding a topic. If you are still having challenges after reading the assigned material twice and working the out the material on paper and the computer, it is your responsibility to seek help. I am here to help and will be glad to assist you in your learning process. Please make an appointment to visit with me on my calendar.

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.

