STT 3860 - Introduction to Statistical Modeling - Spring 2023

Contents

Instructor: Dr. Alan T. Arnholt
Office: Walker Hall 237
Office Hours: M-R 9:30 am – 10:30 am, and by appointment

- Make an appointment to see me by clicking here.

Course Description:

Introduction to Statistical Modeling: Introduction to Statistical Modeling is a continuation of Introduction to Data Management and Visualization with an emphasis on statistical modeling and reproducible reporting using professional tools. Hypothesis testing is introduced via resampling, and bootstrapping is used to introduce estimation. Cross-validation is used to evaluate and select models that take into account the bias-variance trade-off. Supervised learning techniques discussed will include but are not limited to linear regression, regression trees, classification trees, and random forests.

Course Objectives:

- Students will learn to use a reproducible work flow for all assignments.
- Students will improve their technology expertise.
- Students will learn to create statistical models using both supervised and unsupervised learning techniques.
- Students will learn to assess models via cross-validation.

Course Text:

- *An Introduction to Statistical Learning with Applications in R* by Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani

Additional References:

- Predictive Model Building
- bookdown: Authoring Books and Technical Documents with R Markdown
- Data Science with R
- *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.

Tentative Schedule and Course Pacing:

Tentative course assignments and due dates are maintained here.

Course Grading:

- 10% of the course grade will come from 13 DataCamp modules.
- 20% of the grade will come from Rmarkdown summaries of the 13 DataCamp modules.
- 45% of the grade will come from 3 individual bookdown LAB assignments.
- 25% of the grade will come from the final **group** presentation.

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 in Figure 1.

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

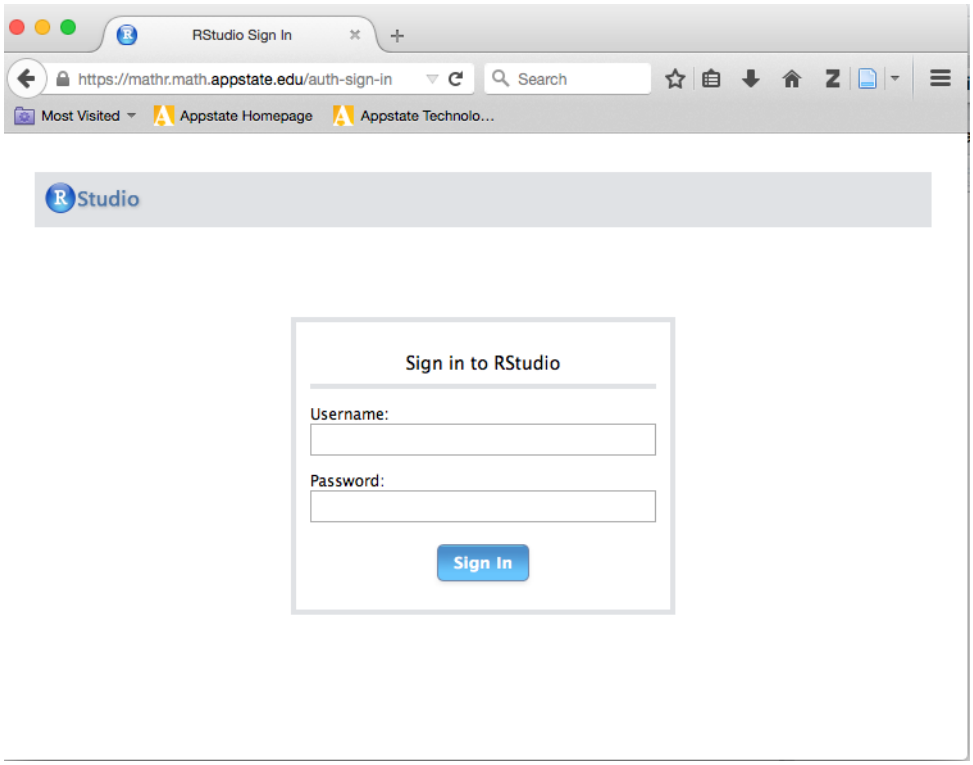

Figure 1: RStudio login screen