STT 5811 Syllabus - Fall 2015

Fall 2015

Instructor: Dr. Alan T. Arnholt

Office: Walker Hall 340

Office Hours: 4:15-5:45 Tuesday, 4:15-5:45 Thursday, 12:15-1:45 Friday

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

Questions related to ISCAM III materials should be addressed with the piazza account for STT 5811. Questions related to *Passion Driven Statistics* materials should be addressed with the piazza account for STT 2810 If you send me a class related email, I will remind you to post questions on the appropriate piazza account.

Course Description:

This course introduces students at the post-calculus level to statistical concepts, applications, and theory. Topics include: comparisons with categorical and numerical data, statistical significance, sampling and sampling distributions, and randomized experiments. Statistical concepts will be developed through simulations, and applications will focus on statistical problem-solving. Students will complete a semester long research project using principles of reproducible research. The semester long research project covers data management; descriptive statistics; inferential statistics, including testing one and two sample procedures; ANOVA; Chi-Square tests; simple regression; and multiple regression. All of the topics will be addressed through a student selected research question that will be answered using one of six large data sets.

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 to create and present graphs for both univariate and multivariate data.
- 5. Students will learn how to construct and test hypotheses using both classical and randomization approaches.
- 6. Students will learn how to construct confidence intervals using both classical and bootstrap approaches.
- 7. Students will learn how to generate random and simple random samples and their relationships to permutation and bootstrap distributions.
- 8. Students will learn how to work with named sampling distributions (t, F, binomial, chi-square, hypergeometric, and normal).
- 9. Students will learn the scope of inferential conclusions for numerous scenarios (experiments, observational studies, etc.).

Course Texts:

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

Investigating Statistical Concepts, Applications, and Methods, Third Edition (ISCAMIII) by Beth L. Chance and Allan J. Rossman

Passion Driven Statistics (Light Weight) (PDS) also available on Rpubs.

Course Grading:

• 30% of the course grade will come from the semester long research project (SLRP). The semester long research project consists of from 9 peer graded assignments submitted via CrowdGrader. 15% of the SLRP will come from your Research Proposal; 15% of the SLRP will be for a poster each student

will create and present. The remaining 20% of the SLRP will be for participation, largely tracked in piazza. The participation tracked in piazza includes questions asked on the forum, questions answered on the forum, etc.

• 30% of the course grade will come from 13 peer graded assignments submitted via CrowdGrader over ISCAMIII investigations/weekly reflections on reproducible research assignments.

NOTE: Grading rubrics and directions for all assignments are provided in the directory CoursePacing of the class repository. Your grade for a peer evaluated assignment will be a combination of peer evaluations and an evaluation of your peer grading. The algorithm used to compute your grade is explained in this presentation.

- 10% of the course grade will be for a midterm exam.
- 10% of the course grade will be for a final exam.
- 10% of the course grade will be for developing/improving PDS assignment rubrics.
- 10% of the course grade will be for participation (largely tracked through piazza). Each time you come to class, you should hand me a 3 by 5 card before sitting down with your first and last name in the top left, the date in year-month-day format (2015-09-03) in the top right, with at least one question you have about the assigned material for that day, which will also contribute to your participation grade.

If you know you will be absent at any point in the semester for any reason, please let me know beforehand; and make sure to submit your work via CrowdGrader before the due date.

How To Get Unstuck

If you have a PDS related question, please ask your question on piazza. All other questions should be addressed on the STT 5811 piazza account. Your classmates may have the same question or may answer your question before I can provide an answer. 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

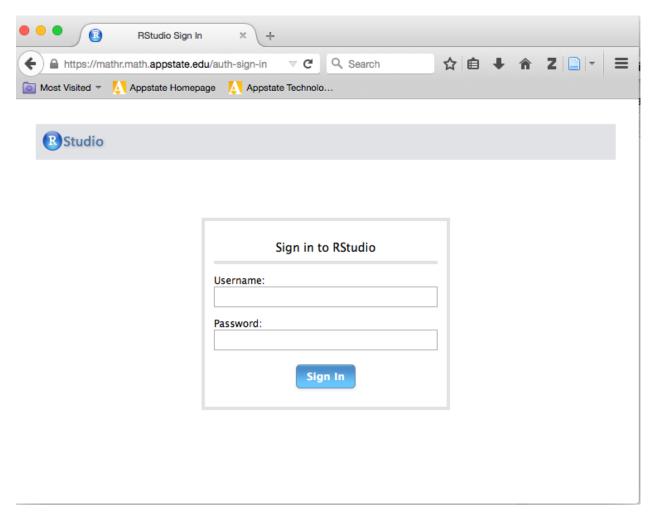
This course conforms with all Appalachian State University policies with respect to academic integrity, disability services, and class attendance. The details of the policies may be found at http://academicaffairs.appstate.edu/resources/syllabi.

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.