STA 4930: Statistical Computing with R
Spring 2022

Instructor: Dr. Justin Shows
Email: jhshows@stat.fsu.edu
OSB 327 TuTh 1:20-2:35
Office: OSB 307
In-Person office hours will be held through 3:00-4:00 Tuesday and Thursday. You can also make an appointment through e-mail. The Zoom link is

https://fsu.zoom.us/j/4722571439 with Meeting ID 472 257 1439.

COURSE OBJECTIVES

Course Description: This course provides an introduction to the R software. Topics include reading and managing data, statistical analyses, graphics, writing functions, simulation studies, and resampling methods. No previous experience with R is required; however, students should be familiar with basic statistical techniques such as regression, analysis of variance, categorical data analysis, and descriptive statistics.

Prerequisite: Instructor's Permission

Note: Since this course strictly utilizes R, it cannot be counted towards the SAS certificate.

MATERIALS

Textbook: None

Software: R software will be used in the course. Previous R experience is not necessary. We will be using the R Studio interface. You will need to download and install both R and R Studio (both free)

R: https://www.r-project.org/ (Choose any CRAN mirror in the US)
R Studio: https://rstudio.com/products/rstudio/download/#download

SYLLABUS CHANGE POLICY

Except for changes that substantially affect implementation of the evaluation (grading) statement, this syllabus is a guide for the course and is subject to change with advance notice.

The syllabus is intended to be an agreement between you (student) and I.

The presence of your name on your section roster after drop/add ends indicates that you have read and understood <u>all</u> of the information and documents posted on the course Canvas site, you have had your questions about them answered, and you accept and will abide by my policies.

ATTENDANCE, ABSENCE AND DROP/WITHDRAWAL

Attendance will be taken the first week of class via an ungraded online quiz. If you are absent, the Registars Office may drop you from the course. I expect you to keep up-to-date with the lessons, activities and announcements by getting in touch with a classmate or with me, **and** by reading the course site on Canvas. Drop and withdrawal deadlines are given in the University's academic calendar at https://registrar.fsu.edu/registration_guide/summer/academic_calendar/

University Attendance Policy

Excused absences include documented illness, deaths in the family and other documented crises, call to active military duty or jury duty, religious holy days, and official University activities. These absences will be accommodated in a way that does not arbitrarily penalize students who have a valid excuse. Consideration will also be given to students whose dependent children experience serious illness.

Academic Honor Policy:
The Florida State University Academic Honor Policy outlines the University's expectations for the integrity of students' academic work, the procedures for resolving alleged violations of those expectations, and the rights and responsibilities of students and faculty members throughout the process. Students are responsible for reading the Academic Honor Policy and for living up to their pledge to "...be honest and truthful and...[to] strive for personal and institutional integrity at Florida State University." (Florida State University Academic Honor Policy, found at http://fda.fsu.edu/academic-resources/academic-integrity-and-grievances/academic-honor-policy.)

Americans With Disabilities Act:
Students with disabilities needing academic accommodation should:
(1) register with and provide documentation to the Office of Accessibility Services; and
(2) bring a letter to the instructor indicating the need for accommodation and what type.
Please note that instructors are not allowed to provide classroom accommodation to a student until appropriate verification from the Office of Accessibility Services has been provided.

This syllabus and other class materials are available in alternative format upon request.
For more information about services available to FSU students with disabilities, contact the:
Office of Accessibility Services
874 Traditions Way
108 Student Services Building
Florida State University
Tallahassee, FL 32306-4167
(850) 644-9566 (voice)
(850) 644-8504 (TDD)
oas@fsu.edu
https://dsst.fsu.edu/oas

ASSESSMENT AND GRADING

- Assignments will be posted on Canvas along with a due date. You may work together on these assignments, but you must formulate and write your own code. You may not share files with other students.
- Only the code from assignments should be submitted through Canvas as one document. This document should be an R script (.R) file or copied and pasted into a .txt file. The file must be compilable.
- The lowest assignment score will be dropped.
- Assignments submitted less than three days after the due date will be penalized by 25%. Assignments submitted more than three days after the due date will not be accepted.
-

Privacy: I strictly follow FERPA rules and guidelines. For example, grades cannot be given out in e-mail. For more information, read https://registrar.fsu.edu/records/ferpa/

Grading:
Assignments   100%
Weighted Average is rounded up to the next integer.

Grading Scale:

Average	Letter
93-100	A
90-92	A-
87-89	B+
83-86	B
80-82	B-
77-79	C+
73-76	C
70-72	C-
67-69	D+
63-66	D
60-62	D-
0-59	F

Introduction to R:

This information mostly comes from the R manual on the R website, and the other websites linked on Canvas. I will not go over everything in the manuals, but will use them as guides.

R is an integrated suite of software facilities for data manipulation, calculation and graphical display. Among other things it has:

- an effective data handling and storage facility,
- a suite of operators for calculations on arrays, in particular matrices,
- a large, coherent, integrated collection of intermediate tools for data analysis,
- graphical facilities for data analysis and display either directly at the computer or on hardcopy, and
- a well developed, simple and effective programming language (called 'S') which includes conditionals, loops, user defined recursive functions and input and output facilities. (Indeed most of the system supplied functions are themselves written in the S language.)

R is commonly used for developing new methods of interactive data analysis. It has developed rapidly, and has been extended by a large collection of packages. However, most programs written in R are essentially written for a single piece of data analysis.

No statistical software is the "best." The preference of a particular software depends on the user and the purpose. There are always advantages and disadvantages of any system. For instance, R is free, while a SAS license is very expensive. SAS can process very large data sets, but sometimes R has problems with this. In many jobs, only one may be available, so it is important to know both. Job prospects typically increase with software knowledge and flexibility. Also knowing Stata, SPSS, Minitab and others is useful.

There is an important difference in philosophy between S (and hence R) and the other main statistical systems. In S a statistical analysis is normally done as a series of steps, and the results are stored in objects. SAS and SPSS will give a great amount of output from an analysis, R will give minimal output and store the results in a fit object for later use by R functions. For example, if you perform regression, R will create an object containing least squares estimates, residuals, fitted values, and more. Very little will be displayed in the console or a window.

There are multiple packages, and each contains a certain set of functions, data sets, and more. The base package (already installed) contains many of the functions for common statistical tasks. To get another package, you must first install (download) it and then load it by using library(package name). We will do this later.

Using Help in R: There is a Help menu on the toolbar. I typically use HTML help, but usually the best thing to do is to perform a Google search. There are many R user groups and threads and someone almost always knows how to do what you want to do. Many professors and instructors also post their course information on websites and allow people to see it.

After R is started, there is a console awaiting for input. At the prompt (>), you can enter numbers and perform calculations. You can also open a script to type your code in, then highlight it and submit it. I prefer to do this so that it can be saved.