STAT 184: Introduction to R

Summer 2023 (Session 2)

Contents

Course Information

Class Time & Location

Sec 201: M-F, 10:10 - 11:10, Forest Resources Building 417

Classes will start from Monday, July 3, 2023 instead of Wednesday, June 28, 2023.

Instructor

Soumya Mukherjee

Office: 333 Thomas Building
Email: szm6510 [at] psu [dot] edu

Office Hours: To be announced shortly (Will be held on Zoom)

Grader

Yuqi Hong (Contact details will be announced shortly)

Website

Canvas: https://canvas.psu.edu

Other relevant links will be updated shortly.

Laptops It is recommended that you bring a computer to class each day if you can, but there's no need or expectation that you buy a computer if you don't already have access to one. You will have access to all the resources you need on the classroom computers. However, it is usually easier to work on your personal machine.

Textbooks

Data Computing (2nd edition) by Daniel Kaplan & Matthew Beckman

Click Here for the Data Computing free ebook.

(*recommended*): **Happy Git and GitHub for the UseR** by Jenny Bryan et al.

Click Here for the Happy Git free ebook.

(*recommended*): **R for Data Science** by Hadley Wickham & Garrett Grolemund

Click Here for the R for Data Science free ebook.

(*recommended*): **Intro to R - tidyverse** by Brendan R. E. Ansell Click here for the Intro to R free ebook.

Other readings assigned as needed.

Software/Tools

Get access to R / RStudio right away! We will be walking through this on the first day of class, but it would be helpful to install these early. Get help if needed (Google, Canvas Discussions, friends, office hours).

- Install BOTH R & RStudio on your computer (requires config). Make sure to download the appropriate version for your operating system (Mac or Windows).
  - R software (free): http://cran.rstudio.com/
  - RStudio Desktop (free) https://rstudio.com/products/rstudio/download/

Make an account on the following websites:

- GitHub: github.com
  - use your student email
  - use your real name
  - make sure your user name is appropriate (you will be using it to turn in assignments)
- Rstudio Cloud: https://posit.cloud/learn/cheat-sheets (use your student email)

IMPORTANT NOTE: If you get stuck in performing any of the above tasks, please do not worry as we will guide everyone through the process on the first day of class and even after that.

Communication

Canvas Discussion Board We will be using canvas Discussion board for all class Q&A, to help you benefit from each other's questions and the collective knowledge of your classmates, instructor, and grader. Questions should be posted to the entire class (for content-related questions). I encourage you to ask questions if you are struggling to understand a concept, and to answer your classmates' questions when you can. Canvas Discussion posts will only be responded to by the Instructor and Grader during business hours.

Do NOT use Canvas Discussion for personal/private matters (grades, accommodations, etc); email those questions or comments to the instructor/grader directly, visit office hours, or discuss them in person after class.

Email Most issues about classroom activities should be posted to Canvas Discussion, but you should use email (or a discuss in person or via Zoom) for all personal or private matters. Emails will only be respond to during business hours. Plan accordingly!

Grading

Learning outcomes will be assessed based on performance in each of the following categories accompanied by their impact on the overall grade:

Pct	Assignment Type
25%	Final Project
30%	Activities (Mini-Projects)
20%	Reading Quizzes
20%	Exercises
5%	In-class Participation

Late Work Nearly all assignments are turned in electronically on Canvas, and most are due at 9:00am (Eastern) on the specified date. The official time stamp on Canvas will generally be used to determine late work.

Late work will be accepted for nearly all homeworks and in class activities. Late work is not permitted at all for some assignments (for e.g., Final project) unless pre-arranged with the instructor.

For other regular assignments (e.g., reading quizzes, homework exercises, & activities):

- 80% credit for the first 24 hours after the due time/date;
- 60% credit between 24 & 48 hours after the due time/date;
- 50% credit after 48 hours beyond the due date/time until the last day of classes, August 9th (not the last day of finals)

Final grades Final grades will be awarded based on the following scale:

Grade	Minimum
A	93%
A-	90%
B+	87%
B	83%
B-	80%
C+	77%
C	70%
D	60%
F	0%

Course Assignments

For most of this course, we will be alternating between lecture material and in class activities. Details will be discussed in class. Exercises and activities will be available on Canvas the day we have the relevant lecture material, then the following class period will be dedicated to working on those activities. Activities are generally due at 9:00 am (before class starts) a day or two later. Always check canvas for exact due dates.

Students are encouraged to start early on problems an ask questions on Canvas Discussions and in office hours. Students are *encouraged* to work together on these assignments, but each student must turn in their own work.

Exam There will be no midterm or final exam as such.

Final Project A successful project will find one or more real & interesting data sets, and use their R programming skills to tell a story that reveals insights from the data. The assigned activities from the Data Computing text book are good examples of the scale and scope of work expected for a successful project, with the differences that you are expected to do the work independently using your own data (not loaded from an R package), and you are responsible for producing the narrative explaining your investigation and conclusions as you work through the analysis.

Course Goals and Objectives

Some goals and objectives may be reduced or expanded as time permits, but a tentative list follows:

- General Tools
  - Become familiar with R programming language
  - Become familiar with RStudio development environment
  - Generate reports and reproducible work with RMarkdown
  - Exposure to Git/GitHub source control
- Navigate basic syntax in R
  - Adopt basic notions of consistent programming style
  - Show proficiency using functions
  - Install and use a variety of contributed R packages
- Read & write multiple data file types using R
  - CSV
  - web scraping
- Show proficiency with basic data wrangling operations using R
  - Principles of "Tidy Data"
  - `tidyverse` package
- Generate descriptive statistics using R
- Show proficiency with layered graphs & data visualization
  - "Glyph-ready Data"
  - `ggplot2` graphics
- Additional topics
  - Basic machine learning
  - Regular expressions
- Possible extensions as time permits
  - Web applications (i.e. Shiny)
  - `lattice` graphics
  - `mosiac` graphics interface
  - Special topics

Policies & Resources

Counseling and Psychological Services (CAPS) Many students at Penn State face personal challenges or have psychological needs that may interfere with interfere with their academic progress, social development, or emotional wellbeing. The university offers a variety of confidential services to help you through difficult times, including individual and group counseling, crisis intervention, consultations, online chats, and mental health screenings. These services are provided by staff who welcome all students and embrace a philosophy respectful of clients' cultural and religious backgrounds, and sensitive to differences in race, ability, gender identity and sexual orientation.

Counseling and Psychological Services at University Park (CAPS):

- Phone: 814-863-0395
- Web: http://studentaffairs.psu.edu/counseling/

Penn State Crisis Line (24 hours/7 days/week): 877-229-6400

Crisis Text Line (24 hours/7 days/week): Text LIONS to 741741

ECoS Code of Mutual Respect The Eberly College of Science Code of Mutual Respect and Cooperation embodies the values that we hope our faculty, staff, and students possess and will endorse to make the Eberly College of Science a place where every individual feels respected and valued, as well as challenged and rewarded.

Academic Integrity Statement Academic dishonesty is not limited to simply cheating on an exam or assignment. The following is quoted directly from the "PSU Faculty Senate Policies for Students" regarding academic integrity and academic dishonesty:

> Academic integrity is the pursuit of scholarly activity free from fraud and deception and is an educational objective of this institution. Academic dishonesty includes, but is not limited to, cheating, plagiarizing, fabricating of information or citations, facilitating acts of academic dishonesty by others, having unauthorized possession of examinations, submitting work of another person or work previously used without informing the instructor, or tampering with the academic work of other students.

All University and Eberly College of Science policies regarding academic integrity/academic dishonesty apply to this course and the students enrolled in this course. Refer to the following URL for further details on the academic integrity policies of the Eberly College of Science: http://www.science.psu.edu/academic/Integrity/index.html. Each student in this course is expected to work entirely on her/his own while taking any exam, to complete assignments on her/his own effort without the assistance of others unless directed otherwise by the instructor, and to abide by University and Eberly College of Science policies about academic integrity and academic dishonesty. Academic dishonesty can result in assignment of "F" by the course instructors or "XF" by Judicial Affairs as the final grade for the student.

Disability Policy Penn State welcomes students with disabilities into the University's educational programs. If you have a disability-related need for reasonable academic adjustments in this course, contact Student Disability Resources (SDR; formerly ODS) at 814-863-1807, 116 Boucke, http://equity.psu.edu/student-disability-resources. In order to receive consideration for course accommodations, you must contact ODS and provide documentation (see the guidelines at http://equity.psu.edu/student-disability-resources/guidelines).

Syllabus Changes

This syllabus is subject to change as circumstances warrant; substantive changes will be distributed in writing (e.g., through Canvas) or announced in class.

This document was last modified on June 26, 2023