MICR 493/593; GEOL 490/590
Reproducible Data Analysis, Fall 2020

University of Tennessee, Knoxville

COURSE FORMAT: Online only; synchronous Wednesdays 3:30 – 4:20 pm
COURSE CREDIT HOURS: 3.0 except with permission of instructor

Faculty Contact Information

Instructor: Dr. Andrew D. Steen (he/him)
Office Hours: by appointment
Office Location: Strong Hall 717, but I am generally at home this semester
Email Address: asteen1@utk.edu
Public calendar: https://goo.gl/woCs32
Course website: https://adsteen.github.io/data_analysis_2020

Course Description/Information:

This course aims to teach introductory principles of reproducible data analysis using the R statistical platform. No prerequisites are required.

Value Proposition:

Students will learn to more efficiently and reproducibly analyze their own data, leading to faster analysis times, deeper analyses, and fewer mistakes.

Student Learning Outcomes/Objectives:

Successful students will demonstrate the following skills:

- Analyzing data using the R programming language
- Sharing and executing code using the Bash shell and Git/GitHub
- Communicating data with clear data visualizations
- Writing clear, easy-to-interpret code by using good coding style
- Creating logical, efficient workflows for data analysis that are generalizable to any programming language
- Producing generalizable software to perform analyses relevant to research

Programmatic Outcomes / Department Goals:

This course will contribute to students' ability to carry out data analysis in the service of scientific research.

Learning Environment:

Class will involve lectures and a heavy component of peer education. Students will work in teams of 3-4 throughout the semester to teach one another.

Course Communications:

I will use the class website https://adsteen.github.io/data_analysis_2020 to post lecture notes, tasks, and other documents. I will also manage communications via Canvas messages.

PLEASE MAKE SURE THAT YOU HAVE CANVAS SET TO DELIVER MESSAGES AT LEAST DAILY. Outside of class hours, email is the best way to contact me. I will seek to respond to email within 24 hours, not counting holidays and weekends.

How to Be Successful in This Course:

STUDENT RESPONSIBILITY:

1. Watch all lecture videos and read all assigned texts.
2. Be present for synchronous classes (unless you have been excused). Participate.
3. Have access to appropriate computing devices for the class. If you don't have what you need, contact Dr. Steen as soon as possible.
4. You will be placed in a peer group of 3-4 students. Meet regularly with this group. Everyone in your group is responsible for everyone else in the group's homework success.
5. Be respectful of students and the instructor.
6. Focus on class activities during class time.
7. Actively contribute to the class' intellectual environment.
8. Abide by the UT Honor Code.

INSTRUCTOR RESPONSIBILITY:

1. Prepare enlightening lectures and synchronous activities.
2. Evaluate students promptly, fairly and equitably.
3. Respectfully challenge students to reach their potential.
4. Maintain a constructive virtual classroom environment. This may include asking disruptive / non-contributing students to leave the classroom.
5. Behave according to University codes of conduct.

Texts/Resources/Materials:

All of these are available for free at the URLs listed.

- Grolemund and Wickham (2017) *R for Data Science*. Available at https://r4ds.had.co.nz/
- Chacon (2014) *Pro Git*. Available for free at https://git-scm.com/book/en/v2
- Healy (2018) *Data Visualization: A Practical Introduction*. Available at https://socviz.co/index.html

Also recommended:

- Wilke (2019) *Fundamentals of Data Visualization*. Available at https://serialmentor.com/dataviz/
- Wickham (2015) *Advanced R*. Available at http://adv-r.had.co.nz/

Required Equipment:

- **Access to an internet-connected computer with the following software installed is essential to success in this class**. If you don't have access to a computer, please talk to Dr. Steen. We can find one for you to use.
- R version 4.0 or later, available at https://cran.r-project.org/
- Rstudio Desktop version 1.3.0 or later, available at https://www.rstudio.com/products/rstudio/#Desktop
- git version 2.28 or later, available at https://git-scm.com/downloads
- Windows 10 computers only: Bash shell. See guide here: https://www.windowscentral.com/how-install-bash-shell-command-line-windows-10
- Windows 8 or below: Bash shell. Probably the best way is to install Cygwin, as described here: https://www.cygwin.com/

Course Resources:
The course Canvas and github sites will be used.

Course Requirements, Assessments, and Evaluations:
Classroom attendance is not required but it is ***strongly*** encouraged.

GRADES WILL BE WEIGHTED AS FOLLOWS:

- Task completion (homeworks): 70%
- Final Project: 30%
  (Including steps along the way to the final project)

GRADES:

A	A-	B+	B	B-	C+	C	D	F
≥93	(93-90]	(90-86]	(86-83]	(83-80]	(80-76]	(76-70]	(70-60]	<60

The instructor reserves the right to curve grades upwards at his discretion. Grades on any assignment may be appealed to the instructor, who may choose to re-grade the entire assignment.

Grading scheme:
Homeworks will largely be peer-graded according to a standard rubric. Completed homeworks will be granted an A. In general, homeworks that are handed in but which do not earn an A will not be accepted. Non-accepted homeworks can be graded if they are turned in by the following Friday; after that they will turn into an F. Homeworks that are not turned in will earn an F. Each student will be granted one 'freebie' homework per class, that can be skipped without consequence. **Importantly, no student in a peer support group will have their homework accepted until all students in the group have turned in their homework**. That is to say, each student will be responsible for all the other students in their group.

Major Assignments and Exams (names and due dates)
The final project is due Tuesday, Nov 24, at 11:59 am.

Course Feedback:
Dr. Steen welcomes direct feedback from students at all time: in class or out of class; in person, via email, or anonymously by unsigned note left in my mailbox. I cannot see this course from a students' perspective, so I rely on your input to teach the best class possible. As with all University of Tennessee classes, students will be asked to provide feedback via SAIS forms.

A note on attendance:
Times are weird. We're all facing different challenges during this pandemic. As a result, I understand that planning is a challenge. Please let me know if you're struggling to make assignments on time, to make it to the synchronous class on Wednesdays, or having any other struggle in the class.

Students will also have the opportunity to meaningfully influence the direction of the class, especially by determining the form of the final project.

Course Topics:

(Provisional schedule, to be updated pending class progress)

Date	Reading due	Homework due
8/19	-	-
8/26	Millman & Perez Economist on Reinhart/Rogoff	HMK 1: What is reproducible research, and why
9/2	Git Ch. 1-4, 6, 8 R4DS Ch 1-2	HMK 2: Scientific computing basics: the shell, Git, and GitHub; RStudio projects, directory structure
9/9	R4DS Ch. 3 MICR 490 R style guide	HMK 3: Basic visualization and input/output
9/16	R4DS Ch 4, 20	HMK 4: Base R: atomic vectors, data types, data frames / tibbles, lists, and subsetting
9/23	R4DS Ch 5, 7	HMK 5: Data transformation and exploration
9/30	Healey Ch 1 - 5	HMK 6: What makes a good plot?
10/7	Wilke chapters relevant to your plot	HMK 7: GOODPLOT/BADPLOT
10/14	R4DS Ch 9 - 16	HMK 8: Wrangling
10/21		Final project prospectus
10/28	R4DS Ch 17 - 20	HMK 9: Programming
11/4	R4DS Ch 21 - 25	HMK 10: Models
11/11	-	-
11/18	-	Final project draft
11/24	-	Final project (due at 11:59 am).

All homework assignments are due at 3:30 pm on Wednesdays. There is no final exam.

Dear Student,

The purpose of this ***Campus Syllabus*** *is to provide you with important information that is common across courses at UT. Please observe the following policies and familiarize yourself with the university resources listed below. At UT, we are committed to providing you with a high quality learning experience. I want to wish you the best for a successful and productive semester.*
Interim Provost John Zomchick

ACADEMIC INTEGRITY

"An essential feature of the University of Tennessee, Knoxville is a commitment to maintaining an atmosphere of intellectual integrity and academic honesty. As a student of the university, I pledge that I will neither knowingly give nor receive any inappropriate assistance in academic work, thus affirming my own personal commitment to honor and integrity."

UNIVERSITY CIVILITY STATEMENT

"Civility is genuine respect and regard for others: politeness, consideration, tact, good manners, gracious-ness, cordiality, affability, amiability and courteous-ness. Civility enhances academic freedom and integrity, and is a prerequisite to the free exchange of ideas and knowledge in the learning community. Our community consists of students, faculty, staff, alumni, and campus visitors. Community members affect each other's well-being and have a shared interest in creating and sustaining an environment where all community members and their points of view are valued and respected. Affirming the value of each member of the university community, the campus asks that all its members adhere to the principles of civility and community adopted by the campus" http://civility.utk.edu/.

EMERGENCY ALERT SYSTEM:

The University of Tennessee is committed to providing a safe environment to learn and work. When you are alerted to an emergency, please take appropriate action. Learn more about what to do in an emergency and sign up for UTAlerts at http://safety.utk.edu. Check the emergency posters near exits and elevators for building specific information. In the event of an emergency, the course schedule and assignments may be subject to change. If changes to graded activities are required, reasonable adjustments will be made, and you will be responsible for meeting revised deadlines.

YOUR ROLE IN IMPROVING TEACHING AND LEARNING THROUGH COURSE ASSESSMENT

At UT, it is our collective responsibility to improve the state of teaching and learning. During the semester you may be requested to assess aspects of this course either during class or at the completion of the class. You are encouraged to respond to these various forms of assessment as a means of continuing to improve the quality of the UT learning experience.

DISABILITIES THAT CONSTRAIN LEARNING:

"Any student who feels he or she may need an accommodation based on the impact of a disability should contact the Office of Disability Services (ODS) at 865-974-6087 in 100 Dunford Hall to document their eligibility for services. ODS will work with students and faculty to coordinate reasonable accommodations for students with documented disabilities." Information on accessibility at UTK is also at http://accessibility.utk.edu

WELLNESS:

The Student Counseling Center is the university's primary facility for personal counseling, psycho-therapy, and psychological outreach and consultation services. http://counselingcenter.utk.edu/ and The Center for Health Education and Wellness engages in prevention and intervention efforts to increase awareness, impact student decision making, and positively influence our university community. The Center manages 974-HELP (also at http://wellness.utk.edu/).