SYLLABUS
Data Wrangling and Management with R
Statistics 295
Spring 2022
Mon/Wed, 3:50–5:10 pm in Lucy Stone Hall A143

Professor: Steve Buyske (Buyske rhymes with "nice key")
buyske@stat.rutgers.edu   (Please include "295" in the subject line)

Canvas: https://rutgers.instructure.com/courses/162988

Office Hours: Monday 2–3 pm (via Zoom while classes are remote),
Thursday 8–9 pm (always via Zoom),
and by arrangement (email me for a time).
Join office hours by using the Zoom link on the Canvas course site.

Please try Piazza, via the link in Canvas, for your questions rather than email.

Required Text: *R for Data Science*
Hadley Wickham & Garrett Grolemund. O'Reilly Media. 978-1491910399
Available at no cost at https://r4ds.had.co.nz
Exercise solutions can be found at
https://jrnold.github.io/r4ds-exercise-solutions/

Prerequisite: 01:198/960:142 or 01:960:212 or 01:960:291 or 01:960:384 or 01:960:379 or 01:960:401

Software: We will be using the free R software, RStudio Cloud, and SQL. No prior experience is required. As long as you have a web browser, no software installation is required.

Note: This course can be used to fulfill the data management requirement of the Data Science certificate and minor. It can also be used in place of 01:960:390 for Statistics majors and minors.

For the first two weeks of classes, we will meet over Zoom on January 19, and then switch to an asynchronous format until the University returns to in-person classes. Details will be explained in the January 19 class and posted in Canvas.

Plan of the course:

- Each week during the remote period there will be new videos, reading, and practice, generally posted on Wednesdays.
- During the in-person period, in addition to the classes there will not generally be videos but there will be reading and practice each week.
- Practice will not be collected or graded—ask questions about it in class, on Piazza, or in office hours.
- There will be a main assignment weekly; the format of the main assignment will vary from week to week.
- During the remote period there will also be small assignments, usually embedded in the videos.
- There will be a final course project, to be described later in the semester.

Assessment:

There will be a main assignment every week other than Week 14. All main assignments will have the same point value. Your two worst scores will be down-weighted to half the weight of a usual main assignment.

<div align="center">

Final Grade Based on the Following
Weekly main assignments	81%
Weekly small assignments	4%
Final project (due May 5)	15%

</div>

Note: If the remote portion of the semester is extended, then the final grade percentage for small assignments will increase by 2 points for each week and the percentage for main assignments will decrease by 2 points for each week.

Up to an additional four points can be earned via active participation on Piazza (you can post question in Piazza anonymously, but you can only get extra credit for non-anonymous posts and replies). You can also earn one additional point for demonstrating that you have the syllabus by emailing me a picture of wild bird by January 25; be sure to include 295 in the subject line.

Pandemic and other crises:

I understand that we are all facing more issues than we would have been a few years ago. If you tell me you are having trouble, I am not going to judge you or think less of you.

As noted above, there are no exams in this course and I will down-weight your two worst main assignments. Furthermore, everyone can complete two main assignments up to 48 hours late with no explanation needed at all. I hope this will take care of most issues. If there is a more substantial issue, just ask and I will work with you.

Academic Integrity:

Cheating undercuts the whole enterprise of education, academics, and science. If I were to cheat in my own work it would end my career. I take cheating very seriously and do not hesitate to refer cases to the Office of Student Conduct. Violations include: cheating, fabrication, plagiarism, denying others access to information or material, and facilitating violations of academic integrity.

The university's policy on Academic Integrity is available at `https://nbprovost.rutgers.edu/academic-integrity-students`

Students will be required to certify the following for each main assignment and for the project: *On my honor, I have neither received nor given any unauthorized assistance on this assignment.*

The following is a rough guide to our topics, but we will adapt it as circumstances require:

    Introduction to the course, to RStudio Cloud, to R, to ggplot, and to R Markdown.

    Data frames, pipes, and the big five verbs: filter, select, mutate, group_by, and summarize.

    Joins in R.

    Tidy data, dates, and reading and writing files.

    Strings and regular expressions.

    Project management and GitHub; map functions.

    Writing functions, missing values, rowwise operations.

    Data cleaning methodology.

    Webscraping.

    APIs, packages, and ggplot themes.

    Tidytext; R and databases.

    Introduction to SQL and relational databases via the R interface.

    Introduction to SQL without R.

    Introduction to NoSQL; odds and ends.

Learning Goals:

Students successfully completing this course should be able to:

- Import, modify, and reshape data from its "wild" form to formats most suitable for data analysis;
- Efficiently clean and modify data;
- Program interactions with databases to efficiently write and receive data.

Relevant Student Services:

Please visit the Rutgers Student Tech Guide, `https://it.rutgers.edu/technology-guide/students/`, for resources available to all students. If you do not have the appropriate technology for financial reasons, please email Dean of Students deanofstudents@echo.rutgers.edu for assistance. If you are facing other financial hardships, please visit the Office of Financial Aid at `https://financialaid.rutgers.edu/`.

Counseling, ADAP & Psychiatric Services (CAPS)
(848) 932-7884 17 Senior Street, New Brunswick, NJ 08901
`http://health.rutgers.edu/medical-counseling-services/counseling/`
CAPS is a University mental health support service that includes counseling, alcohol and other drug assistance, and psychiatric services staffed by a team of professionals within Rutgers Health services to support students? efforts to succeed at Rutgers University. CAPS offers a variety of services that include: individual therapy, group therapy and workshops, crisis intervention, referral to specialists in the community, and consultation and collaboration with campus partners.
Crisis Intervention
`http://health.rutgers.edu/medical-counseling-services/counseling/crisis-interventi`
Report a Concern
`http://health.rutgers.edu/do-something-to-help/`

Violence Prevention & Victim Assistance (VPVA)
(848) 932-1181 3 Bartlett Street, New Brunswick, NJ 08901 `http://www.vpva.rutgers.edu/`
The Office for Violence Prevention and Victim Assistance provides confidential crisis intervention, counseling and advocacy for victims of sexual and relationship violence and stalking to students, staff and faculty. To reach staff during office hours when the university is open or to reach an advocate after hours, call 848-932-1181.

Disability Services
(848) 445-6800 Lucy Stone Hall, Suite A145, Livingston Campus, 54 Joyce Kilmer Avenue, Piscataway, NJ 08854 `https://ods.rutgers.edu/`
Rutgers University welcomes students with disabilities into all of the University's educational programs. In order to receive consideration for reasonable accommodations, a student with a disability must contact the appropriate disability services office at the campus where you are officially enrolled, participate in an intake interview, and provide documentation:
`https://ods.rutgers.edu/students/documentation-guidelines`. If the documentation supports your request for reasonable accommodations, your campus's disability services office will provide you with a Letter of Accommodations. Please share this letter with your instructors and discuss the accommodations with them as early in your courses as possible. To begin this process, please follow the registration process on the ODS web site at: `https://ods.rutgers.edu/students/getting-registered`.