THE OHIO STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

SYLLABUS: STAT 5730
INTRODUCTION TO R FOR DATA SCIENCE
SPRING 2021

Course overview

Instructor

Instructor: Sebastian Kurtek
Email address: kurtek.1@stat.osu.edu
Phone number: 614 292 0463
Office hours: Wednesdays 2PM-3PM, Fridays 2PM-3PM via Zoom (link to office hours will be posted on Carmen)
Office Location: Cockins Hall 440B

Course description

R is a freely available statistical computing environment and programming language. It has become a dominant workhorse for modern statistical research and data analysis, and is being widely adopted in industrial data analytics as well. The primary goal of the course is to teach students how to use R for data analysis: both (1) efficient use of the R computing environment and (2) effective programming in the R language.

There are formal prerequisites for the course. This is a **statistics course**, so the examples and applications demonstrated in the course will be oriented towards data analysis and statistical endeavors. Basic numeracy and familiarity with statistics is expected for motivation and perspective.

Note: This course will be administered 100% online.

Course learning outcomes

By the end of this course, students should successfully be able to:

- Produce dynamic and reproducible reports with R Markdown;

- Visualize various types of data in R using the ggplot2 package;
- Import, manipulate and summarize data in R;
- Use the R tidyr package to clean data prior to statistical analysis;
- Write and execute R functions that involve iterations or conditional statements;
- Build interactive apps via the shiny package in R.

Course materials

Required

1. (R4DS) Grolemund and Wickham (2016): *R for Data Science*. Electronic version: http://r4ds.had.co.nz. This web version of the book can be accessed freely from any web browser.

2. (HoPR) Grolemund (2014): *Hands-On Programming with R*. Electronic version: https://rstudio-education.github.io/hopr/. This web version of the book can be accessed freely from any web browser.

Course technology

For help with your password, university e-mail, Carmen, or any other technology issues, questions, or requests, contact the OSU IT Service Desk. Standard support hours are available at https://ocio.osu.edu/help/hours, and support for urgent issues is available 24x7.

- **Self-Service and Chat support:** http://ocio.osu.edu/selfservice
- **Phone:** 614-688-HELP (4357)
- **Email:** 8help@osu.edu
- **TDD:** 614-688-8743

Baseline technical skills necessary for online courses

- Basic computer and web-browsing skills
- Navigating Carmen

Technology skills necessary for this specific course

- CarmenZoom audio and video meetings
- Recording a slide presentation with audio narration
- Recording, editing, and uploading video

Necessary equipment

- Computer: current Mac (OS X) or PC (Windows 8+) with high-speed internet connection
- Webcam: built-in or external webcam, fully installed

- Microphone: built-in laptop or tablet mic or external microphone

Necessary software

Students will be required to use RStudio software. RStudio can be downloaded for free at https://www.rstudio.com/. **Before downloading RStudio, you must also download and install the base R software at http://www.r-project.org.** You are expected to install R and RStudio on your personal computer by downloading the software from the links above. You will also have to install the packages tidyverse and rmarkdown in R. Further instructions will be provided during the lectures.

Grading and faculty response

Grades

Category	Percentage
Participation	**10%**
Homework	**30%**
Midterm Assessment	**30%**
Final Course Project	**30%**
Total	**100%**

See course schedule, below, for due dates

Assignment information

1. *Participation*: You are expected to watch all lectures. The lectures will be delivered asynchronously and posted on the Carmen course webpage. In addition, there will be discussion boards on Carmen that you are **required** to contribute to regularly.

2. *Homework*: Homework will be assigned (approximately) weekly, will be due on announced dates and will be graded. Learning to compute and program requires practice. Homework assignments will mainly consist of exercises designed to reinforce the concepts covered in class during the previous week. You will be allowed to drop one homework score from your grade. I recommend completing all of the homeworks, even if you plan to drop one.

3. *Midterm Assessment*: A midterm assessment will take place approximately halfway through the course. Details will be announced on Carmen and during one of the lectures.

4. *Final Project*: Students will work individually on a final project involving the manipulation, exploration, and presentation of data. Details will be announced on Carmen and during one of the lectures.

Late assignments

If you absolutely need turn in an assignment late and have a valid excuse, please see me for the necessary arrangements. However, you must notify me in advance in such a situation. Exceptions to this policy will be permitted only in extreme situations such as serious injury immediately prior to an assignment being due or severe illness requiring hospitalization.

Grading scale

93–100: A
90–92.9: A-
87–89.9: B+
83–86.9: B
80–82.9: B-
77–79.9: C+
73–76.9: C
70 –72.9: C-
67 –69.9: D+
60 –66.9: D
Below 60: E

Faculty feedback and response time

I am providing the following list to give you an idea of my intended availability throughout the course. (Remember that you can call **614-688-HELP** at any time if you have a technical problem.)

Grading and feedback

For large weekly assignments, you can generally expect feedback within **14 days**.

E-mail

I will reply to e-mails within **24 hours on school days**.

Discussion board

I will check and reply to messages in the discussion boards every **48 hours on school days**.

Attendance, participation, and discussions

Student participation requirements

Because this is a distance-education course, your attendance is based on your online activity and participation. The following is a summary of everyone's expected participation:

- **Logging in**: **MULTIPLE TIMES PER WEEK**
  Be sure you are logging in to the course in Carmen each week, including weeks with holidays or weeks with minimal online course activity. (During most weeks you will probably log in many times.) If you have a situation that might cause you to miss an entire week of class, discuss it with me *as soon as possible.*
- **Posted lectures**: **REQUIRED**
  You are required to watch all lectures posted on Carmen.
- **Office hours**: **OPTIONAL**
  Attendance for regularly scheduled office hours is optional. If you are required to discuss an assignment with me, please contact me at the beginning of the week if you need a time outside of my scheduled office hours.
- **Participating in discussion forums**: **SEVERAL TIMES PER SEMESTER**
  As participation, you are required to post at least five times during the course of the semester as part of our substantive class discussion on the week's topics. A discussion forum will be posted every week starting in Week 2.

Discussion and communication guidelines

The following are my expectations for how we should communicate as a class. Above all, please remember to be respectful and thoughtful.

- **Writing style**: While there is no need to participate in class discussions as if you were writing a research paper, you should remember to write using good grammar, spelling, and punctuation. Informality (including an occasional emoticon) is fine for non-academic topics.
- **Tone and civility**: Let's maintain a supportive learning community where everyone feels safe and where people can disagree amicably. Remember that sarcasm doesn't always come across online.
- **Citing your sources**: When we have academic discussions, please cite your sources to back up what you say. (For the textbook or other course materials, list at least the title and page numbers. For online sources, include a link.)
- **Backing up your work**: Consider composing your academic posts in a word processor, where you can save your work, and then copying into the Carmen discussion.

Other course policies

Student academic services

Student academic services offered on the OSU main campus
http://advising.osu.edu/welcome.shtml.

Student support services

Student support services offered on the OSU main campus http://ssc.osu.edu.

Academic integrity policy

Policies for this online course

Please help us maintain an academic environment of mutual respect, fair treatment, and personal growth. You are required to produce original and independent work for all assessments. Although students are often encouraged to work together on homework assignments, **all students must submit their own work**.

Ohio State's academic integrity policy

It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term "academic misconduct" includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct http://studentlife.osu.edu/csc/.

Copyright disclaimer

The materials used in connection with this course may be subject to copyright protection and are only for the use of students officially enrolled in the course for the educational purposes associated with the course. Copyright law must be considered before copying, retaining, or disseminating materials outside of the course.

Statement on title IX

Title IX makes it clear that violence and harassment based on sex and gender are Civil Rights offenses subject to the same kinds of accountability and the same kinds of support applied to offenses against other protected categories (e.g., race). If you or someone you know has been

sexually harassed or assaulted, you may find the appropriate resources at http://titleix.osu.edu or by contacting the Ohio State Title IX Coordinator, Kellie Brennan, at titleix@osu.edu

Accessibility accommodations for students with disabilities

The University strives to make all learning experiences as accessible as possible. If you anticipate or experience academic barriers based on your disability (including mental health, chronic or temporary medical conditions), please let me know immediately so that we can privately discuss options. To establish reasonable accommodations, I may request that you register with Student Life Disability Services. After registration, make arrangements with me as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion. SLDS contact information: slds@osu.edu; 614-292-3307; slds.osu.edu; 098 Baker Hall, 113 W. 12th Avenue.

Accessibility of course technology

This online course requires use of Carmen (Ohio State's learning management system) and other online communication and multimedia tools. If you need additional services to use these technologies, please request accommodations with your instructor.

- Carmen (Canvas) accessibility
- Streaming audio and video
- Synchronous course tools

Your mental health!

As a student you may experience a range of issues that can cause barriers to learning, such as strained relationships, increased anxiety, alcohol/drug problems, feeling down, difficulty concentrating and/or lack of motivation. These mental health concerns or stressful events may lead to diminished academic performance or reduce a student's ability to participate in daily activities. The Ohio State University offers services to assist you with addressing these and other concerns you may be experiencing. If you or someone you know are suffering from any of the aforementioned conditions, you can learn more about the broad range of confidential mental health services available on campus via the Office of Student Life's Counseling and Consultation Service (CCS) by visiting ccs.osu.edu or calling 614- 292-5766. CCS is located on the 4th Floor of the Younkin Success Center and 10th Floor of Lincoln Tower. You can reach an on call counselor when CCS is closed at 614-292-5766 and 24 hour emergency help is also available through the 24/7 National Suicide Prevention Hotline at 1-800-273- TALK or at suicidepreventionlifeline.org

Course schedule (tentative)

Week	Dates	Topics, Readings, Assignments Due
1	1/11-1/15	Topic: Introduction to R, RStudio and R Markdown Reading: R4DS 1.4-1.6, 4, 27.1-27.4.3; HoPR 1-2
2	1/25-1/29	Topic: Visualization in R with ggplot2 Reading: R4DS 3, 28
3	2/1-2/5	Topic: Data types and representation Reading: HoPR 1, 3-5 Due: Homework 1
4	2/8-2/12	Topic: Data frames and data manipulation Reading: HoPR 3-4, R4DS 5.1-5.5, 10 Due: Homework 2
5	2/15-2/19	Topic: Data summarization Reading: R4DS 5.6-5.7 Due: Homework 3
6	2/22-2/26	Topic: Data import and workflows Reading: R4DS 6-8, 11.1-11.2, 11.6 Due: Homework 4
7	3/1-3/5	Midterm Assessment
8	3/8-3/12	Topic: Tidy data and relational data Reading: R4DS 12-13 Due: Homework 5
9	3/15-3/19	Topic: Functions and conditional execution Reading: R4DS 19, HoPR 6-7 Due: Homework 6
10	3/22-3/26	Topic: Dates and times Reading: R4DS 16 Due: Homework 7

11	**3/29-4/2**	**Topic: Iteration and functional programming** **Reading: R4DS 21, HoPR 9** **Due: Homework 8**
12	**4/5-4/9**	**Topic: Dashboards and R Shiny** **Reading: R4DS 25** **Due: Homework 9**
13	**4/12-4/16**	**Additional Topics Time Permitting: Strings and factors, Working with many models and list-columns**
14	**4/19-4/23**	**Project Time**