SYLLABUS: STAT 6730
INTRODUCTION TO COMPUTATIONAL STATISTICS
AUTUMN 2020

Course overview

Instructor

Instructor: Dr. Sebastian Kurtek
Email address: kurtek.1@stat.osu.edu
Phone number: 614-292-0463 (contact via e-mail is highly preferred)
Virtual office hours via CarmenZoom: Tuesdays 2PM-3PM (Zoom link posted in Carmen)
Office location: Cockins Hall 440B

Teaching Assistant

Mr. Dalton Hopper, Virtual TA office hour via CarmenZoom: Thursdays 11:30AM-12:30PM
(Zoom link posted in Carmen)

Course description

Computational statistics is an area within statistics that encompasses computational and
graphical approaches to solving statistical problems. Students will learn how to manipulate
data, design and perform simple Monte Carlo experiments, and be able to use resampling
methods such as the Bootstrap. They will be introduced to technologies that are useful for
statistical computing. Through creating customized graphical and numerical summaries
students will be able to discuss the results obtained from their analyses. The topics of the
course include:

1. Introduction to R
2. Dynamic and reproducible reports with R Markdown
3. Data manipulation in R
4. Visualization of data

5. Smoothing and density estimation
6. Generating random variables
7. Monte Carlo simulation
8. The Bootstrap
9. Permutation methods
10. Cross-validation

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;
- Write and execute R functions that involve iterations or conditional statements;
- Apply common computational methods including Monte Carlo, smoothing and density estimation, the bootstrap and permutation methods.

Prerequisites

This course is intended to be taken during the second year of the MAS program. It is expected that students will have exposure at a mathematical level to foundational concepts in probability and statistics including random variables, estimation, hypothesis testing, and linear regression. The formal prerequisites for this course are: STAT 6301 and 6302 or equivalent; STAT 6410 and 6450, or STAT 6910 and 6950, or permission of the instructor. Previous programming experience is not required, but familiarity with computer systems is expected.

Course materials

The primary resource for reading will be slides and additional references assigned for reading by the instructor. There is one required book for the course that will be used for the parts of course dealing with data manipulation and visualization in R. The title of the book and various ways to access it are provided below.

Required

Wickham, H. & Grolemund, G.: *R for Data Science*. O'Reilly Media.

– Web version: http://r4ds.had.co.nz

– Print version: http://shop.oreilly.com/product/0636920034407.do

– Online access to print version: http://proxy.lib.ohio-state.edu/login?url=http:

//proquest.safaribooksonline.com/?uiCode=ohlink&xmlId=9781491910382

Note that the web and print versions have different chapter numbering.
I will list reading for both versions. You can also access the print version online via Safari Books using your OSU login.

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 this course

- Basic computer and web-browsing skills
- Navigating Carmen
- CarmenZoom
- Recording a slide presentation with audio narration
- Recording, editing, and uploading video

Necessary equipment

- Computer: current Mac (OS X) or PC (Windows 10+) 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

- This class requires you to use the statistical software package called R (The R Project for Statistical Computing; http://www.r-project.org/). This software package is available as Free Software.
  - You can download R for Windows, Mac, and Linux, from the CRAN archive at https://cran.r-project.org.
  - An in-depth introduction to R is available at http://cran.r-project.org/doc/manuals/R-intro.pdf
  - Hands-on tutorials are available in the Swirl system, which you can learn about at http://swirlstats.com/. In particular, "R Programming: The basics of

> programming in R" is an appropriate first tutorial for students who have never used R.

- An easier to use interface to R is available in the software package RStudio. This package is available for Windows, Mac, and Linux and can be downloaded for free from http://rstudio.org. **Note that RStudio requires R to be installed.**

Course delivery

- Each week, we will meet once in-person on Monday. The in-person class meetings are intended for delivery of lecture material. The instructor will post accompanying materials on the class website (Carmen). You are responsible for reviewing all materials.

- During most weeks, a laboratory activity using the R statistical software will be posted on Carmen. The laboratory activities will be accompanied by live CarmenZoom interactions.

- All assignments will be posted on the class website. You will be given ample time to complete the assignments.

- I will hold weekly office hours via CarmenZoom. The dates and times will be announced later and posted on the class website.

Grading and faculty response

Grades

Assignment or category	Percentage
Homework	20
Labs (Wednesday Online Meetings)	20
Midterm Project	30
Final Project	30
Total	100

See course schedule, below, for tentative due dates

All course grades will be recorded on the class website (Carmen).

Assignment information

Homework

- Homeworks will generally be assigned on a biweekly basis.
- Students are required to use R Markdown for their homeworks. They should be written in a style that smoothly integrates prose, code, tables and graphics. It should be easily readable.
- You are required to submit both the source code and a hard copy of the generated document in Carmen.
- All homework will be graded on a 3 point scale: 1 point for good-faith effort, 1 point for technically-correct working solutions, 1 point for well-formatted and easily-readable code.

Labs

- Labs will consist of exercises. They will take place during our Wednesday meetings on CarmenZoom.
- Like homework, students are required to use R Markdown and the lab report should be written in a style that smoothly integrates prose, code, tables and graphics. It should be easily readable.
- Submit the source code (Rmd) to Carmen within 24 hours of the end of the class.
- Labs will be graded on the same 3 point scale as the homework.

Projects

- There will be two project assignments. The first project will be due approximately midway through the course and the second will be due on the day of the final exam (there will be no final exam for this course).
- Like homework and labs, students are required to use R Markdown to produce their final project reports. The report should be written in a style that smoothly integrates prose, code, tables and graphics. It should be easily readable.

Homework and lab assignments in the wrong format (e.g., .pdf, plain text, .doc) will receive 0 points automatically, no exceptions.

Late assignments

Generally, late homework and labs will not be accepted. However, if there are extenuating circumstances beyond your control, please contact the course instructor immediately.

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 homework and lab assignments, you can generally expect feedback within **7 days**.

E-mail

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

Attendance and participation

Student participation requirements

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

- **Logging in**: **AT LEAST ONCE 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.*

- **In-person class meetings**: **EXPECTED**

  You are expected to attend all in-person lectures and you are responsible for all material presented during these lectures. However, formal attendance will not be taken during the class.

- **Online labs**: **EXPECTED**

  You are expected to attend and participate in all online labs. You are responsible for completing the weekly lab assignments on time even if you must miss the lab session. Formal attendance will not be taken during the online meeting.

- **Office hours**: **OPTIONAL**

  My office hours will be held on CarmenZoom and are fully 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.

Communication guidelines

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

- **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.)

Other course policies

Health and safety

The Ohio State University Wexner Medical Center's Cornavirus Outbreak site (https://wexnermedical.osu.edu/features/coronavirus) includes the latest information about COVID-19 as well as guidance for students, faculty and staff. Guidelines and requirements for campus safety from the University's COVID-19 Transition Task Force were published on July 1 on the Safe and Healthy website (https://safeandhealthy.osu.edu).

Potential disruptions to instruction

- As much as is possible, students will have access to material online if they are unable to attend class because of positive diagnosis, symptoms, or quarantine required following contact tracing.

- If the instructor is unable to be present in person because of positive diagnosis, symptoms, or quarantine following contact tracing, a new instructor will be assigned to the course. Details will be given on the course website if this arises.

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 hybrid course

- **Homework, lab and project assignments**: You are expected to produce original and independent work for lab and project assignments. Although students are often encouraged to work together on homework assignments, all students must submit their own written work **in their own words**. Note that allowing others to copy your work is considered academic misconduct. Academic misconduct will not be tolerated and will be dealt with procedurally in accordance with University Rule 3335-31-02. (This policy can be found at http://oaa.osu.edu/coam.html.)

- **Reusing past work**: In general, you are prohibited in university courses from turning in work from a past class to your current class, even if you modify it. If you want to build on past research or revisit a topic you've explored in previous courses, please discuss the situation with me.

- **Falsifying research or results**: All research you will conduct in this course is intended to be a learning experience; you should never feel tempted to make your results or your library research look more successful than it was.

- **Collaboration and informal peer-review**: The course includes many opportunities for formal collaboration with your classmates. While study groups and peer-review of major written projects is encouraged, remember that comparing answers on an assignment is not permitted. If you're unsure about a particular situation, please feel free just to ask ahead of time.

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 (Recommended)

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. In light of the current pandemic, students seeking to request COVID-related accommodations may do so through the university's request process, managed by Student Life Disability Services. 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; http://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

Disclaimer

This syllabus should be taken as a fairly reliable guide for the course content. However, you cannot claim any rights from it and in particular we reserve the right to change due dates or the methods of grading and/or assessment if necessary. Any changes will be communicated to you through official course announcements.

Course schedule (tentative)

This schedule is subject to revision. Students are expected to attend class meetings and to regularly check for updates.

Week	Dates	Topics, Readings, Assignments, Deadlines
1	W 8/26	**Course Overview, Introduction to R** **Reading: R4DS (web: 4, print: 2)**
1	M 8/31	**Data Types in R** **Reading: R4DS (web: 6, print: 5)**
2	W 9/2	**Lists and data frames in R** **Reading: R4DS (web 10; print 7)**
2	M 9/7	**No Class**
3	W 9/9	**Data manipulation and summarization in R** **Reading: R4DS (web 5; print 3)**
3	M 9/14	**Grammar of graphics** **Reading: R4DS (web 3; print 1)**
4	W 9/16	**Lab 1: ggplot2** **Due: Homework 1**
4	M 9/21	**Smoothing**
5	W 9/23	**Lab 2: Smoothing**
5	M 9/28	**Functions** **Reading: R4DS (web 19; print 15)**
6	W 9/30	**Lab 3: Functional programming**

		Due: Homework 2
6	M 10/5	**Density estimation**
7	W 10/7	**Midterm Project Help**
7	M 10/12	**Generating random variables**
8	W 10/14	**Lab 4: Conditionals and iteration** **Reading: R4DS (web 21–21.4; print 17, pp 313–324)** **Due: Midterm Project**
8	M 10/19	**Monte Carlo simulation**
9	W 10/21	**Lab 5: Monte Carlo** **Due: Homework 3**
9	M 10/26	**Split, apply, combine**
10	W 10/28	**Lab 6: plyr**
10	M 11/2	**The Bootstrap**
11	W 11/4	**Lab 7: The Bootstrap**
11	M 11/9	**Pipes and dplyr** **Reading: R4DS (web 5, 18; print 3, 14)** **Due: Homework 4**
12	W 11/11	**No Class**
12	M 11/16	**Cross-validation**
13	W 11/18	**Lab 8: Pipes and dplyr** **Reading: R4DS (web 21.4–21.9, 22; print 17, pp322–344, 18)**
13	M 11/23	**Permutation Tests**
14	W 11/25	**Lab 9: Cross-validation**
14	M 11/30	**TBD** **Due: Homework 5**
15	W 12/2	**Final Project Help**
Exams	R 12/10 4PM	**Final Project Due**