Math 410B, Winter 2021
Advanced Statistical Methods (Part 2)

Instructor: Dr. Kathy Temple
Office: Samuelson 221D
Phone: 963-1389
E-mail: Kathryn.Temple@cwu.edu
Office Hours: MW 10 - 10:50 AM, TTh 1:00 - 1:50 PM, and by appointment

Meeting times: For the first two weeks, plan on being on Zoom from 9 - 9:50 M - Th. After the first two weeks, if university policy and public health conditions allow, you can choose to attend in person. We'll plan on MW 9 – 9:50 in Samuelson 251; TTh 9 – 9:50 in Samuelson 138 Computer Lab (unless otherwise announced).

Course Prerequisites: Math 410A, or permission of the instructor.

Course Goals and Description: This course is the first of a two-part sequence designed to cover the most common statistical methods in use today. We will learn the theory behind the models and methods we use, and we will spend a lot of time using them to answer questions about real-world data. Topics in the first quarter will likely include linear regression, multiple regression, logistic regression, inference from regression, interaction terms, linear discriminant analysis, variable selection methods, and one- and two-way ANOVA.
The second half (410B) will include many details about ANOVA, non-parametric statistics, permutation and bootstrap methods, ridge regression, the lasso, time series analysis, cluster analysis, and Principal Component Analysis. This part of the course will also include a significant component on presenting statistics in public.
Students completing both parts A and B of Math 410 have seen almost all the material on the Society of Actuaries' *Statistics for Risk Modeling* exam.

Required Text: Gareth James, et al. *An Introduction to Statistical Learning with Applications in R*, Springer. (Note that this text is freely available for download.)

Evaluation and Grading: Grades will consist of the following components: (approximately) biweekly labs (15%), homework (20%), a quarter-long "R Journal" (5%), two take-home exams (15% each), a public speaking component (15%), and a final project (15%), in which you wil complete and write about a significant statistical analysis of your (possibly plural) own.
Grades will be assigned according to the following scale:

		A	93-100%	A- 90-92.99%
B+	87-89.99%	B	83-86.99%	B- 80-82.99%
C+	77-79.99%	C	73-76.99%	C- 70-72.99%
D+	67-69.99%	D	63-66.99%	D- 60-62.99%
		F	59.99% and below	

Details concerning course components

Lab Assignments

Labs will be assigned approximately biweekly. Lab due dates will be given with the assignment, but will often be 5PM on Thursdays. Lab write-ups must be typed, written in complete sentences, and follow the general guidelines for written work given below.

We will be making use of the statistical software R in class, and you will be using R for labs and exams. R is a free, open-source software package that is incredibly powerful. I strongly recommend the interface R Studio and this is what I will be using and demonstrating in class. R has a bit of a

learning curve, but it's well worth working through! R is becoming more and more of a standard in industry and other disciplines.

Homework

A small amount of homework will be assigned from the book most weeks. These are designed to let you practice the basics of the new statistics we are covering, and the corresponding techniques in R. Due dates will be announced at the time the homework is assigned, but will typically be 5PM on Mondays. Since we're shifting to slightly less frequent lab assignments this quarter, you may be handing in more 'lab-type' work ask homework.

Exams

There will be two take-home exams this quarter, one around midterms and one near the end of the term.

R Journal

At the end of this course, I want you to have an R script that will include all of the key ideas/techniques from the course along with comments that describe how to use the code. This is the kind of thing that can be super-useful to you later! Because, as we all know, procrastination is a thing, you'll be required to submit your script up to that point every other Thursday by 5pm, with the last one due during the last week of classes. If you want to extend your R Journal from last quarter, feel free! You can also start a new file if you would prefer that.

Public Speaking

This is one of the few places where we explicitly include public speaking skills in our curriculum. This is a crucial skill, and one that can't be fully taught in one class. Rather, the class will provide you with some tools which will hopefully improve your speaking during the quarter, but also enable you to continue to improve after the course is done. We will be completing a series of public speaking assignments. Several of these will involve a short reading, watching some videos or talks, and giving a thoughtful analysis of what you observed. You will also prepare slides for a presentation.

After completing these steps, each person (or team) will give a talk about their work in class. Talks will be recorded, and students will take their videos home, watch them, and reflect and comment on their performance.

You are encouraged to take your work on the final project and consider giving a talk at SOURCE (the Symposium of University Research and Creative Expression). This is an excellent opportunity, and one not to be missed. Abstracts are due near the very beginning of spring quarter.

Final Project

Either by yourself or as part of a team (your choice) you will be responsible for finding a (rich) data set, completing an exploration and analysis of that data, presenting that analysis to the class (see Public Speaking, above), and submitting a written report. More details about this to come!

General Course Policies: Come to class. Do your own work. Work really hard; this class is likely to be quite difficult, but you will leave it with a useful set of skills, and a better understanding of statistics!

Face Coverings and Social Distancing: Due to COVID-19, and under the directive and mandate of public health officials and the president of Central Washington University, students must adopt face covering protocol before entering any classroom or building at CWU until further notice. Students must also follow the social distancing placement marks in buildings and classrooms. If you do not have a face covering Central Washington University can provide one for you. If you have not yet received your CWU-supplied facial covering, please go the SURC Information Desk. Please do so prior to the start of your first class. Face coverings must cover both the mouth and nose. Your mask protects me; my mask protects you. Masks with one-way valves for exhalation don't have the protective value for others - please don't use them as your face covering for any

in-person components. Thank you for helping keep all of us safe!

Class Recordings: Lecture/discussion and any lecture-like portions of lab will be recorded and posted on Canvas for those who can't attend synchronously on a given day. Please plan on attending synchronously on an ongoing basis, however. We all benefit from the conversations in-class, and they will be richer for having you there!

   If you are attending virtually, please turn on your camera if that is reasonable given your situation, bandwidth issues, etc. If you cannot turn on your camera, please use a profile picture on Zoom that matches your profile picture on Canvas (and please put one up on Canvas) so that we don't have to look at a sea of names/initials!

Written Work: All work handed in for the course must be written neatly, legibly, clearly, using correct mathematical notation, and with sufficient explanation. A good rule of thumb is to write your solution so that a classmate who knows roughly what's going on in the course but doesn't know how to do this particular problem can understand your solution. As a side benefit, this makes it much more likely that you will be able to understand your solution when you go back to study for exams! The bottom line: for any written work handed in for the course, including lab write-ups, exams, and homework, *you must show your work*.

Late Work: No late homework will be accepted, but the lowest homework score will be dropped, because life happens. For labs and the take-home exams, assignments can be submitted late with a 20% penalty per 24-hour period, up until the time that graded labs/exams are returned to the rest of the class. *Because we have fewer labs this quarter, there are no dropped labs.* If you have extenuating circumstances around the exams, I encourage you to communicate with me as soon as possible; see the section on communication, below.

Submitting Work Electronically: All work for this course will be submitted electronically through Canvas.

- Labs write-ups and take-home exams are required to be typed and may be in .doc, .docx, or .pdf formats only. If you are using an alternate word processing program, please make sure that you know how to generate one of these formats.

- R scripts are to be submitted as either RStudio scripts (.r files) or RMarkdown files (.rmd). They may not be submitted as text files. This applies to code submitted with labs and take-home exams and to your R Journal.

- If you choose to type your homework, it should be in one of the formats for lab write-ups, above.

- If you choose to hand-write your homework, you have a couple of options for submission:
    - You can scan it to a PDF document using a cell phone camera. Some apps that do this (there are many others if you don't like these): Adobe (free) or CamScanner (free version) for Android phones; the Notes App (built-in) or CamScanner(free version) for iPhone.
    - You can take a well-lit, easy-to-read photograph, insert it into a Microsoft Word or OpenOffice Writer document, make sure that the photo is oriented correctly and easy to read, and then save the document as a .doc, .docx, or .pdf file for upload to Canvas.
    - If your work is multiple pages, please upload it as a SINGLE file.

   Whichever method you choose, please double-check that your image is oriented correctly. Images not oriented correctly or images that are not easy to read will lead to the assignment being returned without being graded.

Academic Honesty: Consult university policies (CWUP 5-90-040(22), CWUR 2-90-040(22), and WAC 106-125-020) for student conduct, cheating, plagiarism, and other academic expectations. CWU's policies and recommendations for academic misconduct will be followed, leading to disciplinary action up to and including failing the course.

Inclusivity Statement: As a member of a peer learning community, a high degree of professionalism is necessary. **CWU expects every member of the university community to contribute to an inclusive and respectful classroom culture.**

Classroom Conduct: Students in this class are expected to interact with students and the professor professionally. Instances of disruptive conduct, obstructive conduct, or harassment (see definitions below from the Washington Administrative Code: WAC 106-125-020) will be referred to the Dean of Student Success.

Disability Support Services: Central Washington University is committed to creating a learning environment that meets the needs of its diverse student body. Students with disabilities should contact Disability Services to discuss a range of options to removing barriers, including accommodations: Hogue Hall 126, 509.963.2214, DS@cwu.edu

Is my absence excused? Excused absences will not lower your overall grade in this class and are determined on a case-by-case basis. Excused absences include illness, bereavement, and school-related activities. With the exception of illness, documentation is required. Excused absences do not include travel for holiday breaks, work, or non-emergency travel delays.

In compliance with RCW 28B.137.010, Central Washington Universitymakes every effort to deal reasonably and fairly with students who, because of religious obligations, have conflicts with scheduled exams, assignments or required attendance. Students must present written notice to their instructor within the first two weeks of class listing the specific dates on which accommodations are required. Contact the Dean of Student Success at (509) 963-1515 for further information or questions.

Communication: This is going to be an unusual quarter in many ways. Some of you are completely remote. Some of us may need to quarantine in the middle of the quarter. We are all starting the quarter completely remote, and we all may need to shift entirely remote on short notice (and it may be worth thinking in advance about how you would make that work, both in terms of technology and in terms of being able to attend class and complete work remotely). In all things, communication is going to be key. If you have something going on, please let me know (no details you don't feel comfortable sharing, of course, just a general heads-up). As things need to change during the quarter, I will do my best to let you know as soon as I can. In all things, please extend grace to me and your classmates, and I will do my best to do the same.

I reserve the right to change the policies contained in this syllabus as dictated by developments during the quarter.

VERY Tentative Schedule

	Class	Lab	Other
Week 1 - 1/5	Introduction and syllabus; ANOVA	ANOVA (one lab day)	Find final project dataset!
Week 2 - 1/11	Cross-validation and the bootstrap	CV and bootstrap (full lab)	R Journal check-in 1.
Week 3 - 1/19	More linear models	Linear models; project check-in 1	Public speaking assignment 1.
Week 4 - 1/25	Ridge/lasso regression	Ridge/Lasso regression (full lab)	Public speaking assignment 2. R Journal check-in 2.
Week 5 - 2/1	Time series!	Project check-in 2	Exam 1 out.
Week 6 - 2/8	More time series!	Time series (full lab)	R Journal check-in 3.
Week 7 - 2/16	Trees	Project check-in 3	Public speaking assignment 3.
Week 8 - 2/22	PCA and cluster analysis	PCA (full lab)	R journal check-in 4.
Week 9 - 3/1	More PCA and cluster analysis	Cluster analysis	Exam 2 out.
Week 10 - 3/8	TBD	TBD	R Journal check-in 5. Public speaking assignment 4.
Finals - 3/15	Final presentations		Final projects due.