THE OHIO STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

SYLLABUS: STAT 3301
STATISTICAL MODELING FOR DISCOVERY I
AUTUMN 2020

Course overview

Instructor

Instructor: Katie Smillie
Email address: smillie.7@osu.edu
Office hours: Virtual office hours via Carmen Zoom

Both I and Professor Hans (the instructor for another section of Stat 3301) will hold several different virtual office hours each week via Carmen Zoom. Students in any section of Stat 3301 may attend any of these office hours. The schedule and Zoom links will be provided on Carmen.

Office Location: Virtual

Due to the COVID-19 pandemic, I do not plan to be in my office regularly. Instead, I will be available virtually over email and Zoom.

Graders

Yongqi Liu (liu.8224@osu.edu)	Jiahao Ping (ping.24@osu.edu)
David Ruttenberg (ruttenberg.6@osu.edu)	Rui Zhang (zhang.9473@osu.edu)

Office hours: Virtual office hours via the MSLC

The graders for the course will hold several virtual office hours during the week via the Mathematics and Statistics Learning Center (MSLC). The hours and instructions for attending these sessions will be provided on Carmen.

Course description

Statistical models for data analysis and discovery, with primary focus on linear regression models. The challenges of building meaningful models for data are explored, and emphasis is

placed on model building and the use of numerical and graphical diagnostics for assessing model fit. Interpretation and communication of the results of analyses is emphasized.

Prerequisite: C- or better in Stat 3202. Prereq or concur: Math 2568.

Course learning outcomes

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

- Formulate regression models that describe relationships between variables and understand the models' statistical foundations.
- Perform a complete regression analysis and communicate the results in both statistical and problem-specific terms.
- Use linear regression methods to build models for large data sets and use the results of the analysis to recommend actions.
- Evaluate and compare different regression models using formal statistical methods and graphical techniques.

Course materials

Required

We will use the textbook *Applied Linear Regression, Fourth Edition* (2014) by Sanford Weisberg. An electronic version of the book can be accessed for free through The Ohio State University Libraries at https://library.ohio-state.edu/record=b8665795~S7. You will need to click on "Connect to resource EBSCOhost"; you may also need to supply your OSU credentials. The online resource is best suited for screen reading; each individual is allowed to print/e-mail/save/download a limited number of pages.

Required reading and homework will be assigned from the book throughout the semester. Be sure to use the fourth edition (red/orange cover) and not the third edition (green cover). The material in the textbook will be supplemented with additional course notes as necessary throughout the semester. More information about the textbook can be found at http://users.stat.umn.edu/~sandy/alr4ed/.

Optional materials

We will occasionally use a graphics package called **ggplot2**. Examples of how to use this package to create statistical graphics will be given in class. The book "ggplot2: Elegant Graphics for Data Analysis" (Second Edition) by Hadley Wickham provides a detailed description of the software. If you are on campus or connected via the library's "off campus" sign in, you should be able to download the book for free at the following link: http://www.springer.com/us/book/9783319242750. I recommend reading the book if you are interested in developing your ggplot2 skills, but you are not required to read the book for Stat 3301.

Another useful resource is the book "R for Data Science" by Garrett Grolemond and Hadley Wickham. The book can be accessed for free at http://r4ds.had.co.nz/. This reference many come in handy if you run into trouble working with data in R.

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

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" provides interactive courses that can be used to review working with R.

- An easy-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.**
- This class requires the use of the (free) R Markdown authoring framework to complete assignments. Information about R Markdown will be provided in class; an online guide with overview information can be found at https://rmarkdown.rstudio.com.

Course delivery

The course will use a mix of **synchronous** and **asynchronous** content.

Required **synchronous** content will be presented live over CarmenZoom on Mondays during the scheduled class time. Students are expected to attend and participate in these live, online class meetings. The synchronous meetings on Mondays will be used to contextualize the previous week's asynchronous content using examples, discussion and questions, with new material introduced as appropriate.

Required **asynchronous** content will be made available each Monday on Carmen. This material will include:

- A 10-15 minute video providing an overview of what you should expect to learn as you work through the rest of the week's asynchronous content.
- Reading assignments from the textbook.
- A short reading quiz that can be completed on Carmen.
- Videos representing the equivalent of two 55-minute lectures that provide in-depth discussion of the topics for the week.
- Homework problems to help assess your understanding of the material.

The instructors for all three sections of Stat 3301 will hold office hours several times during the week via CarmenZoom. The schedule and Zoom links will be posted on Carmen. Students in any section can attend any of these office hours.

The graders for the course will hold virtual office hours several times during the week through the Mathematics and Statistics Learning Center (MSLC). The schedule and information for how to attend these virtual sessions will be made available on Carmen.

Grading and faculty response

Grades

Assignment or category	Percentage
Reading Quizzes	**15**
Homework	**20**
Midterm 1	**20**
Midterm 2	**20**
Final Exam	**25**
Total	**100**

Reading quizzes based on the assigned reading will be posted weekly on Mondays on Carmen and will be due by 11:59pm each Wednesday. These short quizzes are meant to help you focus on important aspects of the reading and to ensure that you have completed the reading before watching the video lectures and attempting the homework. The reading quizzes will be auto-graded in Carmen for accuracy. I will drop your lowest reading quiz score (which can be a quiz that was not completed by the deadline) when computing your reading quiz grade for the semester.

Homework will be assigned approximately weekly. While adjustments may need to be made during the semester, I expect that homework assignments will be due on Carmen on Thursdays by 11:59pm. You will use R Markdown to prepare your homework solutions; instructions for this will be given at the beginning of the semester. Guidelines for preparing and submitting your solutions on Carmen will be provided with the homework assignments.

Typically, homework must be turned in by the deadline to receive credit. If exceptional circumstances (sudden onset of illness, unexpected family situations, etc.) arise, contact the instructor (smillie.7@osu.edu), to discuss the possibility of an extension. We plan to be flexible for those facing unexpected difficulties, though the likelihood of approval of this accommodation may decrease for a student who consistently turns in homework late.

I will drop your lowest homework score (which could be an assignment that was never turned in) when calculating your final homework grade for the semester.

Homework will be graded by the course graders. If your homework does not earn full marks, the grader will leave basic feedback on Carmen pointing you to the trouble spots in your answers; however, due to the size of the course, the grader will not be able to leave detailed feedback. To make up for this, I will provide detailed solutions to the assignments to help you understand how the question could be approached. Comparing your answers to the provided solutions is a good way to self-assess your understanding of the material and to prepare for the exams. Questions about grading should be directed to the graders during their MSLC office hours.

Exams: There will be two midterm exams administered during the semester to assess your understanding of the course material as the semester progresses. The first midterm will be on Wednesday, September 30th and the second will be on Wednesday, November 4th. The final exam will be held during exam week. The exams will be made available to you online, and you will have a window of time in which to complete the exam and upload your answers. The time window will be scheduled well in advance of the exam, and we will work with students to ensure the avoidance of conflicts.

You will be able to use all of your course materials as a reference during the exams. You must complete all exams on your own without assistance from anyone other than a course instructor. You may not provide assistance to anyone on the exam. You may not communicate about the exam with anyone until the exam window for all students has closed.

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 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 **7 days**.

E-mail

I will reply to e-mails within **24 hours on school days**. Specific technical questions about the course material that require significant back-and-forth communication are not well suited for e-mail; while I will do my best to answer such questions, I may ask that you attend one of the virtual office hours that will be spread out during the week if your question isn't easily answerable over email.

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:

- **Attending online, synchronous class meetings**: **ONCE PER WEEK**
  Students are expected to attend and participate in the online, synchronous Monday class meetings. I anticipate that, for some students, circumstances related to COVID-19 might preclude attendance at times during the semester. As much as possible, please stay in contact with the instructor so that we can discuss appropriate accommodations should they be needed.
- **Logging in**: **AT LEAST TWO ADDITIONAL TIMES PER WEEK**
  Be sure you are logging in to the course in Carmen each week, including weeks with holidays. You will need to log in to Carmen to complete reading quizzes, view video content and upload homework assignments. (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*.
- **Office hours**: **OPTIONAL OR FLEXIBLE**
  All office hours are optional. If you need to speak with me privately about a topic that cannot be easily discussed during office hours, please contact me to schedule a time to meet.

Other course policies

Health and safety

The Ohio State University Wexner Medical Center's Coronavirus 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

- The instructor will work to accommodate any student who is unable to participate in class for a period of time because of a positive diagnosis, symptoms, or quarantine required following contact tracing related to COVID-19.

- A back-up instructor has been assigned to this course. In the event the principal instructor is unable to participate in class activities due to a positive diagnosis, symptoms, or quarantine required following contact tracing related to COVID-19, the back-up instructor will take over course duties until the principal instructor is able to resume instruction.

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

- **Reading Quizzes, Midterms and Final Exams**: You must complete all reading quizzes, midterms and final exams on your own without assistance from anyone other than a course instructor. You may not provide assistance to anyone on a quiz or exam. You may not communicate about a quiz or exam with anyone until the quiz deadline has passed or until after the exam window for all students has closed. You will be able to use all of your course materials as a reference during the quizzes and exams.
- **Homework**: You may work together on the homework, but do not copy any part of your solutions from another person or another source. While study groups are allowed, remember that you must produce your own, original work. If you're unsure about a particular situation, please feel free 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. **No course materials provided by the instructor (notes, videos, recordings, computer code, homework assignments, homework solutions, quizzes, exams, etc.) may be distributed publicly or privately to anyone outside of the class.**

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. 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)

The following tentative course schedule is subject to change. The schedule for each week will be posted on Carmen on Mondays.

Week	Dates	Topics, Readings, Assignments, Deadlines
1	Aug 25-28	Introduction, summarizing data in R
2	Aug 31–Sep 4	Review of one-sample problems, relationships between variables
3	Sep 8-11	Intro to simple linear regression (SLR), parameter estimation
4	Sep 14-18	Inference and prediction under SLR models
5	Sep 21-25	Sources of variability, coefficient of determination, residuals
6	Sep 28–Oct 2	Transformations of regressors and response variables
7	Oct 5-9	Multiple linear regression: estimation, inference and prediction

8	Oct 12-16	Interpreting coefficients, correlated predictors
9	Oct 19-23	Regression with factor variables
10	Oct 26-30	Polynomial and flexible mean models
11	Nov 2-6	Multiple categorical regressors, comparing nested models
12	Nov 9-10, 12-13	Regression model diagnostics
13	Nov 16-20	Confounding, causation and missing data
14	Nov 23-25	Common violations of independence
15	Nov 30–Dec 4	Approaches to model selection