



THE OHIO STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

SYLLABUS: STAT 3301
STATISTICAL MODELING FOR DISCOVERY I
AUTUMN 2022
M W F 10:20-11:15 AM POMERENE 150

Course overview

Instructor

Instructor: Daryl Swartzentruber
Email address: Swartzentruber.20@osu.edu
Office hours (Cockins Hall 217):
Monday 12:30-1:30
Wednesday 9:00-10:00
Friday 11:30-12:30

Graders

Grader: Zhenbang Jiao (jiao.180@osu.edu)

Course description

Statistical models for data analysis in the linear regression framework. The challenges of developing meaningful models for data are explored, with emphasis on the model building process, the use of numerical and graphical diagnostics for assessing model fit, and interpretation and communication of results. Statistical foundations are introduced along with basic inferential techniques.

Prerequisite: C- or better in 3202; or 4202 and 5730; or permission of the instructor. Prereq or concur: Math 2568; or permission of the instructor.

Course delivery

This course meets **Mondays, Wednesdays, and Fridays in Pomerene 150 from 10:20 - 11:15 AM**. Lectures will be delivered in person during the scheduled class meeting times. Students are expected to attend and participate in these in-person class meetings. Class meetings will be used to provide in-depth investigation of the topics for the week using a lecture format. Students will participate in these class sessions by engaging in discussions prompted by the instructor and by asking and answering questions. Students should plan to take notes during class.

Carmen

This class will use Carmen. In Carmen, you will find copies of the syllabus, homework assignments, lecture notes and other important documents. Carmen will also be used to keep track of your assignment grades. Additionally, materials for lectures will be uploaded to Carmen.

Course learning outcomes

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

- Use graphical and numerical summaries of data to describe relationships between variables.
- Formulate, fit, evaluate and compare regression models that describe relationships between variables.
- Understand and be able to describe the statistical foundations of standard regression models.
- Identify common violations of the assumptions that underly standard regression models.
- Perform a complete regression analysis and communicate the results in both statistical and problem-specific terms.
- Distinguish between descriptive and causal interpretations of regression.

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.

Students are also required to register with **TopHat**, which is free for Ohio State University students. Please go to the TopHat home page (<https://tophat.com/>) and either login (<https://app.tophat.com/login>) or signup for an account (<https://app.tophat.com/register/>). Make sure you are using the latest version of the app or are accessing it through a browser. TopHat is already connected to the CarmenCanvas course, so it should be listed as one of your courses when you sign in. If not, you can use the join code **861042** Please contact the instructor if you have difficulty accessing TopHat. You are expected to bring a device to class (phone, tablet, laptop, etc.) that can access TopHat.

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

Technology skills necessary for this specific course

- Basic computer and web-browsing skills
- Navigating Carmen

Necessary equipment

- Computer: current Mac (OS X) or PC (Windows 10+) with high-speed internet connection

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 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>.
- During class we will use

Grading and faculty response

Grades

Assignment or category	Percentage
TopHat	15
Homework	20
Midterm 1	15
Midterm 2	15
Data Analysis Project	15
Final Exam	20
Total	100

Tophat questions will be given during class periods. Typically about three questions will be given per class. You will receive 0.8 points for submitting an answer and 0.6 additional points for getting the answer correct. Your total score will be your number of points divided by the number of questions. Thus if you answer all the questions you only need to get a third of them correct to get full credit. Alternatively if you answer about 80% of the questions but get all of those correct you will get full credit. Typically, no opportunities will be given to make up those questions if you miss class.

Homework will be assigned approximately weekly and will be due on Carmen. Homework problems that require R software should be completed in R Markdown and a knitted html file should be uploaded. Homework problems that do not require R may be handwritten (electronically, or on paper and scanned) and uploaded. **You are encouraged to work together on homework; however, each student must produce their own assignment to be handed in. Do not copy any part of another student's homework.** 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.

Exams: There will be two midterms and one final exam administered during the semester to assess your understanding of the course material as the semester progresses. The final exam will take place at the time and date established by the University. Information about the exams will be posted well in advance on Carmen. The dates of the exams are given below.

Midterm Exam 1: Friday October 05, 2022 (during class)

Midterm Exam 2: Wednesday November 04, 2022 (during class)

Final Exam: Wednesday, December 14 (10:00-11:45 AM)

Data Analysis Assignment: There will be an individual, comprehensive data analysis project that will be completed in parts throughout the semester. The data analysis will have three components, with due dates spread throughout the semester:

- Part I: Exploratory data analysis (due mid-October)
- Part II: Preliminary model building and analysis (due mid-November)
- Part III: Model expansion, checking and summarization (due by the end of the semester)

Information about each of the three components will be given during the semester. Each part will count for 5% of the final course grade so that the entire data analysis project counts for 15% of the final course grade.

Late assignments

Generally, late homework assignments are not accepted. Please plan your time so that you can complete assignments far enough in advance to avoid any last-minute problems uploading your completed work; last minute requests for extensions will generally not be granted. If exceptional circumstances (sudden onset of illness, unexpected family situations, etc.) arise, contact the instructor to discuss the possibility of an extension.

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** 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 office hours if your question isn't easily answerable over email.

Attendance, participation, and discussions

Student participation requirements

The following is a summary of everyone's expected participation:

- **Attending online, synchronous class meetings: THREE TIMES PER WEEK**
Students are expected to attend and participate in the in-person class meetings.
- **Logging in: AT LEAST ONCE 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 view the reading assignments and lecture 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.

I expect that you will read and follow the guidelines and requirements for campus safety, which are available at <https://safeandhealthy.osu.edu>.

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

- **Exams:** You must complete the midterm and final exams yourself, without any external help or communication.
- **Written assignments:** Your written assignments should be your own original work. In formal assignments, you should cite the ideas and words of your research sources.
- **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.
- **Collaboration and informal peer-review:** While study groups and peer-review of major written projects are permitted, remember that comparing answers on a quiz or assignment is not permitted. If you're unsure about a particular situation, please feel free to ask the instructor.

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

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

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

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
1	Aug 23-26	Introduction, summarizing data in R
2	Aug 29–Sep 2	Relationships between variables
3	Sep 6-9	Intro to simple linear regression (SLR), parameter estimation
4	Sep 12-16	Inference and prediction under SLR models
5	Sep 19-23	Sources of variability, coefficient of determination, residuals
6	Sep 26–30	Transformations of regressors and response variables
7	Oct 3-7	Multiple linear regression: introduction, estimation
8	Oct 10-14	Interpreting coefficients; inference and prediction
9	Oct 17-21	Correlated predictors, regression with categorical predictors
10	Oct 24-28	Categorical and continuous regressors
11	Oct 31-Nov 4	Interaction effects 1, multiple categorical predictors
12	Nov 7-10	Interaction effects 2, nested models, non-linear regression
13	Nov 14-18	Residuals, leverage, outliers, and influence
14	Nov 21-22	Common violations of independence
15	Nov 28–Dec 2	Causal inference, model comparison, cross validation
16	Dec 5-7	Model selection: all subsets and stepwise regression