



THE OHIO STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

SYLLABUS: STAT 3301

STATISTICAL MODELING FOR DISCOVERY I

AUTUMN 2023

Course overview

Instructor

Instructor: Dr. Matthew Pratola

Email address: mpratola@stat.osu.edu

Office hours: Virtual Office Hours via Carmen Zoom.

Mondays: 2:00pm–3:00pm and 4:00pm–5:00pm

Graders

Contact information: provided on Carmen.

Office hours: Online (information available on Carmen).

Class Meetings

MWF from 10:20am–11:15am in-person Mendenhall Lab (ML) 0191.

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.

Prerequisite or concurrent: Math 2568; or permission of the instructor.

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=b7651844~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.

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 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 be delivered in-person. You are expected to attend lectures and participate in class. Mondays will be used to introduce new material and provide context for the content using examples, discussion and questions. I will also provide an overview of what you should expect to learn as we work through the rest of the week's content.

Each week's content will be made available each Monday on Carmen. This material will include:

- Reading assignments from the textbook.
- Slides that provide in-depth discussion of the topics for the week. The slides will typically have unfilled portions that we will fill out together during class lecture as we explore the topic being discussed.

- Homework problems to help assess your understanding of the material.

I will hold office hours on Monday afternoons. The schedule and Zoom links will be posted on Carmen.

The graders for the course will hold virtual office hours several times during the week. 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
Homework	20
Midterm 1	20
Midterm 2	20
Data Analysis Assignment	15
Final Exam	25
Total	100

Homework will be assigned approximately weekly on Mondays and will be due on Carmen the following week on Wednesdays by 10:00am. Homework should be completed in R Markdown and a knitted html file should be uploaded to Carmen (information on this will be provided at the start of the semester). **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 to assess your understanding of the course material as the semester progresses. The first midterm will be on Wednesday September 27th and the second will be on Wednesday November 8th. The final exam will be held on Tuesday December 12th. All exams will be taken remotely as “take-home” (un-

proctored) exams. You will have a 24-hour window during which you can start the exam; once started, you must complete the exam and upload your solutions in a specified amount of time.

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 assist anyone on the exam. You may not communicate about the exam with anyone until the exam window for all students has closed.

Data Analysis Assignment: A data analysis assignment will be due near the end of the semester. Students will receive a data set that can be used to answer specific questions about a general problem area that will be described in the assignment. The goal is to use the methods and techniques learned in Stat 3301 to help answer the questions posed in the description of the problem. Students will complete the assignment by using R Markdown to prepare a written report that includes appropriate graphics and statistical analysis. More detailed information about the assignment and guidelines for preparing the report will be provided midway through the semester. There will be no homework due during the week or two leading up to the due date for the data analysis assignment.

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 virtual 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:

- **Attend class meetings:**
Students are expected to attend and participate in class meetings.
- **Logging in:**
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 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

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 all midterm and final exams on your own without assistance from anyone other than the course instructor. You may not assist anyone on an exam. You may not communicate about an exam with anyone until after the exam window for all students has closed. You will be able to use all your course materials as a reference during the exams.
- **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

Religious Accommodations

It is Ohio State's policy to reasonably accommodate the sincerely held religious beliefs and practices of all students. The policy permits a student to be absent for up to three days each academic semester for reasons of faith or religious or spiritual belief.

Students planning to use religious beliefs or practices accommodations for course requirements must inform the instructor in writing no later than 14 days after the course begins. The instructor is then responsible for scheduling an alternative time and date for the course requirement, which may be before or after the original time and date of the course requirement. These alternative accommodations will remain confidential. It is the student's responsibility to ensure that all course assignments are completed.

Accessibility accommodations for students with disabilities

The university strives to make all learning experiences as accessible as possible. In light of the recent 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, Readings, Assignments, Deadlines
1	Aug 23-25	Introduction, summarizing data in R
2	Aug 28–Sep 1	Relationships between variables
3	Sep 6-8	Intro to simple linear regression (SLR), parameter estimation
4	Sep 11-15	Inference and prediction under SLR models
5	Sep 18-22	Sources of variability, coefficient of determination, residuals
6	Sep 25–Sep 29	Transformations of regressors and response variables
7	Oct 2-6	Multiple linear regression: introduction, estimation
8	Oct 9-11	Interpreting coefficients; inference and prediction
9	Oct 16-20	Correlated predictors, regression with categorical predictors
10	Oct 23-27	Categorical and continuous regressors
11	Oct 30-Nov 3	Interaction effects 1, multiple categorical predictors
12	Nov 6-8	Interaction effects 2, nested models, non-linear regression
13	Nov 13-17	Residuals, leverage, outliers, and influence
14	Nov 20	Common violations of independence
15	Nov 27–Dec 1	Causal inference, model comparison, cross validation
16	Dec 4-6	Model selection: all subsets and stepwise regression

Syllabus version

July 16, 2022: original

Aug 22, 2022: tentative midterm dates added

Aug 25, 2022: added final exam date

Aug 13, 2023: revised for AU2023 in-person