



THE OHIO STATE UNIVERSITY

---

COLLEGE OF ARTS AND SCIENCES

# **SYLLABUS: STAT 6450**

## **APPLIED REGRESSION ANALYSIS**

### **SPRING 2023**

## **Course overview**

### **Instructor**

Instructor: Subhadeep Paul

Email address: [paul.963@osu.edu](mailto:paul.963@osu.edu)

Lectures: **In person Wednesdays and Fridays, 3:00–4:50 pm at Baker Systems 260.**

Office hours: **In person at Cockins Hall 231 (Wednesdays 10:30-11:30 AM)**

### **Teaching Assistant**

Yi Tang Chen ([chen.8414@osu.edu](mailto:chen.8414@osu.edu))

## **Course description**

Statistics 6450 is intended to be an introduction to regression analysis techniques. Its focus will be on the application of linear regression models in practice but will also cover basic theory of the linear model. Topics of Stat 6450 include:

Simple Linear Regression (SLR) model

- Methodology for fitting models
- Statistical inference
- Diagnostics for verification of assumptions
- Transformations, weighted least squares

### Multiple Linear Regression (MLR) model

- Methodology for fitting models and use of matrix algebra
- Statistical inference
- Binary indicator (1/0) and qualitative predictors
- Piecewise linear models
- Diagnostic measures of model fit

### Variable selection and model building

- Stepwise regression methods
- Lasso and ridge regression
- Bias-variance tradeoff and cross validation

### Generalized Linear Models (GLMs)

- Logistic regression – methodology for fitting, statistical inference, diagnostics, model selection, prediction, ROC curves.
- Poisson regression (if time permits)

## **Prerequisite or corequisite:**

Statistics 6201 or equivalent.

## **Course learning outcomes**

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

- Understand the motivation of regression analysis
- Understand the theoretical assumptions behind the linear model and their importance in properly conducting a regression analysis
- Know how to estimate the parameters in regression models
- Be able to validate the modeling assumptions with formal tests and visual diagnostic tools
- Know how to make inferences regarding the linear model
- Be able to build and validate regression models in a principled manner
- Be able to apply the above knowledge and techniques in on your own data or problems

## Course materials

### Required

The required textbook for this course is:

*Applied Linear Regression*, 4th Edition by Sanford Weisberg. E-book freely available from the Library (unlimited users):

at <https://library.ohio-state.edu/record=b8665795~S7>

### Optional materials

A related book we will also follow in lectures (not required)

*Applied Linear Regression Models*, 5th edition, by Kutner, Nachtsheim, and Neter. (print)

References for learning R programming:

*R programming for Data Science* by Roger D Peng

[https://bookdown.org/rdpeng/rprogdatascience/Links to an external site.](https://bookdown.org/rdpeng/rprogdatascience/Links%20to%20an%20external%20site)

*R for Data Science* by Garrett Grolemund and Hadley Wickham

<https://r4ds.had.co.nz/>

## 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](mailto:8help@osu.edu)
- **TDD:** 614-688-8743

### Technology skills necessary for this specific course

- Zoom
- Writing R codes for statistical analysis and communicating results
- uploading files in carmen.

## 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>
- 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.**
- [Microsoft Office 365 ProPlus](#) All Ohio State students are now eligible for free Microsoft Office 365 ProPlus through Microsoft's Student Advantage program. Each student can install Office on five PCs or Macs, five tablets (Windows, iPad® and Android™) and five phones.
  - Students are able to access Word, Excel, PowerPoint, Outlook and other programs, depending on platform. Users will also receive 1 TB of OneDrive for Business storage.
  - Office 365 is installed within your BuckeyeMail account. Full instructions for downloading and installation can be found <https://ocio.osu.edu/kb04733>.

## Course delivery

The course will be primarily delivered **in person** (Please see this for details on the in-person delivery mode: <https://keeplearning.osu.edu/understanding-instruction-modes>). Occasionally we will have recorded, or synchronous online class delivered via CarmenZoom.

Each week we will cover approximately 220 minutes of content in total. You will be responsible for attending in person (or occasional recorded or live online lectures) as well as studying the material that is assigned. You will be given ample time to complete the assignments.

The instructor will hold weekly office hours at Cockins Halls 231. The times are given above.

# Grading and faculty response

## Grades

Assignment or category	Percentage
Homework	35
Midterm Exam	20
Final Exam	25
Final Project	20
Total	100

## Assignment information

**Homework:** Homework will be assigned regularly (a total of 5 homework assignments). It will consist of both written problems and computer programming/data analysis problems. You are encouraged to work together on the problems, but each student must hand in their own work, written in their own words. Do not copy any part of another student's homework including computer output. Use of homework solutions distributed in previous offerings of the course or available on the web constitutes academic misconduct and will be handled according to university rules. **All homework must be submitted online as a PDF file through the class website (carmen).** Please be sure that the questions are clearly labeled, all supporting work (including computer code) can be easily identified, and that all figures/tables are referenced and interpreted in the text.

**Please note late submission of assignments will not be accepted unless prior exception has been sought.** if you are unable to complete an assignment on time, please get in touch with me as soon as possible so we can discuss your situation

**Exams:** There will be two exams --**one midterm exam and a final exam.** Coverage includes lecture material, assigned reading, and homework. All exams are closed book/closed notes and will be **delivered through carmen quiz and proctored online through CarmenZoom.** Further details will be given in advance of each exam.

**Tentative Date for midterm exam: Wednesday March 1 during class time.**  
**Final Exam Date (University scheduled): Thursday April 27, 4:00-5:45 PM.**

Statistical tables will be provided as needed. Calculators may be used, but no communication devices are allowed (e.g. mobile phones). Makeup exams require a valid excuse and official proof if I am notified in advance or as soon as possible.

**Project:** A data-analysis group class project will consist of an oral presentation of results obtained from analyzing a dataset, and will require use of the R software. The students are expected to work in groups of 3-4 and each group can select a dataset of their choice to analyze (preferably from the research/educational domain of the students to bring more diversity of topics). No written report is required.

**The groups will present their results to the class on the last day of the course (Apr 21) during the usual class time. Slides will be due the day before. More information will be provided through carmen.**

## 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 large weekly assignments, you can generally expect feedback within **14 days**.

### E-mail

I will reply to e-mails related to course logistics within **72 hours on school days**. Please visit the office hours for more technical and content related questions.

# Attendance, participation, and discussions

Students are expected to regularly participate, discuss, and answer questions in in-person lectures. However, students may miss class, for a variety of reasons, including those related to COVID-19 pandemic. As much as possible, please stay in contact with the instructor so that we can discuss accommodations should they be needed.

## 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

- 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 alternative arrangements will be made. Details will be given on the course website

### 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 course

- **Exams:** You must complete the midterm and final exams yourself, without any external help or communication.
- **Written assignments:** Your written assignments, including discussion posts, should be your own original work. You are allowed to discuss with your peers and work together on the assignments. However, the submission must be written in your own words
- **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.
- **Group projects:** This course includes group projects, which can be stressful for students when it comes to dividing work, taking credit, and receiving grades and feedback. I will make the guidelines for group work as clear as possible for each activity and assignment, but please let me know if you have any questions.

### 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

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 at [titleix@osu.edu](mailto: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](mailto:slds@osu.edu); 614-292-3307; <http://slds.osu.edu>; 098 Baker Hall, 113 W. 12th Avenue.

## **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](http://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](http://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)

Week	Dates	Topics, Readings, Assignments, Deadlines
1	Jan 11, 13	Introduction, background materials, data sources
2	Jan 18, 20	Simple Linear Regression (SLR) – parameter estimation and inference
3	Jan 25, 27	SLR – inference on model and ANOVA,
4	Feb 1, 3	Visual diagnostics, lack of fit test, data analysis
5	Feb 8, 10	SLR- diagnostic tools, remedies, transformations, simultaneous inference, data analysis
6	Feb 15, 17	Recap of Linear Algebra, SLR in matrix form, Multiple Linear Regression (MLR) –introduction and estimation, models with interactions
7	Feb 22, 24	MLR – inference, diagnostics and remedy, general linear test
8	Mar 1, 3	<b>Midterm Exam (Mar 1)</b> , Weighted least squares, indicator (dummy) and qualitative predictors
9	Mar 8, 10	Piecewise linear models, polynomial regression, bootstrap
10	Mar 15, 17	<b>Spring Break: No classes</b>
10	Mar 22, 24	Variable Selection and Model building (stepwise regression)
11	Mar 29, 31	Bias variance tradeoff, cross validation, lasso, ridge
12	Apr 5, 7	Data Analysis with MLR, Generalized Linear Models (GLM) – logistic regression estimation, inference
13	Apr 12, 14	Logistic regression – residuals, diagnostics, model selection, prediction, Data Analysis
15	Apr 19, 21	Poisson regression, <b>project presentation (Apr 21)</b>
16	Thursday, Apr 27	<b>Final Exam, Thursday April 27, 4:00-5:45 PM</b>