

## Applied Regression Analysis (Fall 2023)

Face to face: STAT4043-62127 for undergraduate credit, STAT5543-65692 for graduate credit

Online: STAT4043-63362 for undergraduate credit, STAT5543-65694 for graduate credit

**Instructor:** Dr. Pratyaydipta Rudra, 301D MSCS, [prudra@okstate.edu](mailto:prudra@okstate.edu)

### Shorter version of the syllabus

**Class Meetings:** Monday, Wednesday, and Friday, 10:30am to 11:20am, MSCS 310.

**Office Hours and contact:** In person: Mondays between 11:30am-12:30pm Virtual: Tuesdays and Thursdays between 2-3pm (for everyone). Tuesdays between 6-7pm (only for online sections). More options and details on office hour policy [here](#). Email is preferred for contacting me. I will usually respond to your email by the end of the next business day.

**Course Description:** Matrix algebra, simple linear regression, residual analysis techniques, multiple linear regression, dummy variables, interactions, model building tools, logistic regression (if time permits).

**Prerequisites:** One of STAT 4013, 4033, 4053, 5013, or equivalent. More details [here](#).

**Textbook and Materials:** *A Second Course in Statistics: Regression Analysis*, 8th (or at least 7th) edition, by Mendenhall and Sincich. Chapters 1-9 will be covered. The textbook is optional. I do not follow the book strictly.

Most assignments will involve computer-assisted modeling. We will use the statistical software R. More details [here](#).

**Online Classroom:** We will use Canvas course management system for distributing course materials, assignments and grading. See [here](#) for help regarding how to use Canvas.

**Attendance:** This is primarily a lecture course, although we will do some computing and practice exercises during our class time. Livestream option via MS Teams will be provided for online sections. The students in face-to-face sections are expected to attend in person. Recorded lectures will also be provided for the online sections. Attendance will be taken in class, but it will not be used for grading purposes.

**Grading:** Final grade will be a weighted average of unannounced short quizzes, homework assignments, two mid-term exams and a final project. Graduate and undergraduate students will have the same weights, but graduate students may have extra work on some assignments.

Short quizzes	15%
Assignments	25%
Midterm Exams	40%
Final Individual Data Analysis Project	20%

None of the homework or exam scores will be dropped and no make-up exams will be taken except extreme circumstances. More details [here](#).

**Final Grade:** A: >90%, B: 75% to 89%, C: 65% to 75%, D: 50% to 65%, F: 0% to 49%. However, I reserve the right to lower the letter grade cutoffs.

**Use of electronics in class:** Computer use is not a problem, but keep electronic device use limited to academic activities. On some days you might want to execute R code as I work through an example on the class computer, and therefore I strongly encourage bringing your laptop in class, if you have one. Please don't use your cell phone in the class.

**Academic Integrity:** You may discuss homework assignments with other students and look through others' computer code to help debug errors, but each person must write and execute computer code individually and submit your own write-ups. You may not copy computer code or homework write up from others. Absolutely no collaboration is allowed for the other types of assessments including the project and short quizzes. Use of AI not permitted. More details [here](#).

**Special Accommodations:** If needed, details are [here](#).

**Scheduling proctored exams (only for online sections):** You must take the exam either with the face-to-face sections, or with a live external proctor following the instructions in the file "How to schedule exam proctors" under the syllabus module. Please complete this as soon as possible to ensure proctor availability.

**Syllabus Attachment:** This includes information about important add/drop dates and frequently asked questions. You will find a copy under your syllabus module.

**COVID-19 or other sickness policy:** Mask is not mandatory but encouraged if you think you may have respiratory illness. Please don't come to class if sick and use the live streaming option instead. Please let the instructor know ASAP if you miss classes due to a verifiable sickness or other important reason in case you miss an exam or in-class quiz and it might be rescheduled for you.

## Detailed version of the syllabus

### Office Hours Policy:

1. There are one in person and two virtual office hour slots for everyone.
2. In person: **Mon between 11:30am-12:30pm** Virtual: **Tue and Thu between 2-3pm (for everyone). Tuesdays between 6-7pm (only for online sections).**
3. MS Teams links provided in the announcement at the beginning of the semester should be used to join the virtual office hours.
4. For virtual office hours, the student needs to let me know at least an hour before the office hour that they are coming to make sure I am there.
5. Beyond scheduled office hours, you can email or speak with me about other times and we will schedule appointments accordingly.

**Expectation of previous knowledge and skills:** A 4000 level STAT class at OSU or equivalent. Ideally, your previous statistics course(s) covered the normal, t, F, and chi-square distributions; hypothesis testing with critical regions and p-values. If you are using a prerequisite override, please know that a student who only took a course similar to STAT 2013, 2023, or 2053 might struggle with some course materials. I will do some linear algebra review sessions, but some prior algebra knowledge is helpful.

We will use the statistical software R (<https://www.r-project.org>) which is free and available for Windows, Macintosh, and Linux. We will use [R-studio](#) for using R, but you can use any other interface for using R if you like. You may not use other statistical software for this class. Prior experience of statistical computing with R, although not required, can be helpful.

**Details on academic integrity:** Oklahoma State University is committed to the maintenance of the highest standards of integrity and ethical conduct of its members. This level of ethical behavior and integrity will be maintained in this course. Participating in a behavior that violates academic integrity (e.g., unauthorized collaboration, plagiarism, multiple submissions, cheating on examinations, fabricating information, helping another person cheat, unauthorized advance access to examinations, altering or destroying the work of others, and fraudulently altering academic records) will result in your being sanctioned. Violations may subject you to disciplinary action including the following: receiving a failing grade on an assignment, examination or course, receiving a notation of a violation of academic integrity on your transcript (F!), and being suspended from the University. You have the right to appeal the charge. Contact the Office of Academic Affairs, 101 Whitehurst, 405-744-5627, <http://academicintegrity.okstate.edu>.

Please see Details of exams and assignment for what level of collaboration is allowed for each type of assessment. Use of AI is not permitted to write answers, but it can be used for learning.

**Special accommodations:** If you have a documented disability that may require assistance, the instructor will work with you and Student Accessibility Services (SAS) to provide reasonable accommodation to ensure that you have a fair opportunity in the class. I MUST receive documentation from SAS in order to provide accommodations, and this documentation must be received in advance of any assignment that would require assistance. Their website: [SAS](#).

**Resources:**

[Online Classroom](#): You will log in using your OKEY credentials.

[Help regarding how to use Canvas](#)

**Details of exams and assignments:**

**Short quizzes:** 6-7 unannounced short quizzes will be conducted. For online students, these will be conducted over Canvas. For face-to-face students, some quizzes will be conducted in-class and some online. These short quizzes (usually takes no more than 15 minutes to complete) are meant to keep you on track in the course. The quizzes conducted online are typically posted on Fridays and are available for the students to take for a week. The quiz description will tell you which particular topics are covered. They consist of short answer type conceptual questions; you won't need computer coding to answer them. You can consult the lecture materials, textbook or any other resources but you may not consult with any other person. Collaboration of any kind will be considered violation of class policy.

There will be an additional quiz due on the first Friday of class that will consist of a syllabus quiz and a first day of class survey. The next weekly quiz will be a pre-requisite quiz. Other quizzes will be based on recently taught class materials.

**Homework Assignments:** Assignments are made of continuous evaluations. There will be 5 graded homework assignments. These are typically posted on Fridays, and are due in a week. Each assignment will not have the same number of points; your final percentage will be based on total points. Graduate students will have one additional assignment. You may discuss homework assignments with other students. I encourage you to try the homework on your own first and then discuss with peers. However, you must write and execute computer code individually and submit your own write-ups. Late homework assignments will ordinarily not be accepted. Homework assignments are due before the beginning of class (i.e., at 10:30am) on the due date (typically Fridays) and can only be submitted online on Canvas.

Besides the graded homework assignments, I will often provide you practice problems to solve at home. The students are responsible to complete them, but may seek help during office hours.

I am happy to check if you are correctly solving them and provide feedback depending on how much time I have. No solution keys are provided for the practice problems.

**Midterm Exams:** At this point I plan to have two closed-book and closed-notes midterm exams in class on Monday, 10/9 and Friday, 11/17.

The first midterm will cover simple linear regression and matrix algebra. It will focus on concept questions, matrix algebra, and interpreting results from R. You may bring a calculator, but only arithmetic computations will be necessary.

Online students may arrange for a 50-minute proctored exam on 10/9 or 10/10. It might also be possible to take the exam with the face to face students in the classroom depending on availability of space in the classroom (contact me if you want to use this option).

The second in-class midterm exam will be covering all materials taught since the first midterm. This will also be closed-book and closed note. You may use calculator.

Online students may arrange for a 50-minute proctored exam on 11/16 or 11/17. It might also be possible to take the exam with the face to face students in the classroom depending on availability of space in the classroom (contact me if you want to use this option).

**Final Project:** Each of you will conduct a multivariate regression analysis on a dataset provided to you. I may also allow you to bring your own dataset if you like, but the dataset must be approved by me. You will propose an appropriate model to answer research questions, analyze the data, and write a short report with your conclusions. More information on project requirements will be given after the first midterm. The project report will be due electronically at 11:59pm on Friday 12/8.

### **Tentative schedule of assignments, exams and the first two quizzes**

The following table is for giving you an idea about the schedule. Please know that these dates might change. I will tell you in class, send emails or set up announcements in Canvas regarding any changes. I will **not** update the changes in this file.

Additional ungraded tasks such as reading assignments and practice problems may be assigned at the end of any class. These are not included here.

Week	Topics, Assignments and exams
Week 1 (8/21 – 8/25)	<p>Introduction to R and Linear algebra</p> <p>Quiz 1 due (syllabus + survey)</p> <p>Quiz 2 posted (Prerequisites)</p>
Week 2 (8/28 – 9/1)	<p>Handling data in R, Correlation analysis</p> <p>Quiz 2 due (Prerequisites)</p> <p>Assignment 1 assigned</p>
<p>Week 3 (9/4 – 9/8)</p> <p>No class on 9/4</p>	<p>Introduction to simple linear regression (SLR)</p> <p>Assignment 1 due</p>
Week 4 (9/11 – 9/15)	<p>Estimation and hypothesis testing in SLR</p> <p>Assignment 2 assigned</p>
Week 5 (9/18 – 9/22)	<p>Linear algebra part II, Matrix notation for regression</p> <p>Assignment 2 Due</p> <p>Assignment 2.5 assigned</p>
Week 6 (9/25 – 9/29)	Interval estimation, Prediction, ANOVA
Week 7 (10/2 – 10/6)	<p>Simulations using R, Regression diagnostics</p> <p>Assignment 2.5 due</p>
Week 8 (10/9 – 10/13)	<p><b>10/9: Midterm Exam 1</b></p> <p>Remedial measures, Regression modeling pitfalls</p>
Week 9 (10/16 – 10/20)	<p>Outliers, Introduction to Multiple linear regression (MLR)</p> <p>Assignment 3 assigned</p>
Week 10 (10/23 – 10/27)	<p>Categorical predictors, Confounding, Extra sum of squares</p> <p>Assignment 3 due</p> <p>Assignment 4 assigned</p>

Week 11 (10/30 – 11/3)	<p>Multicollinearity, Interactions</p> <p>Assignment 4 due</p> <p>Assignment 5 assigned</p>
Week 12 (11/6 – 11/10)	<p>Polynomial Regression, Introduction to Logistic regression</p> <p>Assignment 5 due</p>
Week 13 (11/13 – 11/17)	<p>Model building</p> <p><b>11/17: Midterm Exam 2</b></p> <p>Data Analysis Project assigned</p>
<p>Week 14 (11/20 – 11/24)</p> <p>Thanksgiving holidays, no class</p>	
Week 15 (11/27 – 12/1)	Model building for a regression analysis problem
Week 16 (12/4 – 12/8)	<p>More examples and review</p> <p>12/8: Data Analysis Project due</p>