Introduction to Regression Analysis (STAT 3220)

Spring 2026

"All models are wrong, but some are useful." –George Box

What is regression and why should you care?

Will UVA win March Madness this year? How much money you will spend on a vacation? What makes a Spotify song "danceable"? Is there really evidence of global warming? What factors affect your chance of being accepted to graduate school? If you have ever asked yourself any of these questions, then you might have tried to use regression! Regression is one of the most fundamental analysis techniques in statistics that uses a set of explanatory variables to try to predict or estimate a certain outcome. Regression, just as the field of statistics in general, can be used in any discipline to answer many interesting questions. In this course, we will explore three foundational techniques of regression: multiple linear regression, logistic regression, and ANOVA.

What will you be able to do?

During this course, you will develop the fundamental practical and interpersonal skills to think and act like a statistician, meaning you value the importance of and are confident in your ability to evaluate evidence.

By the end of this course you will be able to:

- Compare and contrast regression techniques in order to decide which is the most appropriate given a particular research question or data set
- Use a statistical analysis to evaluate a research claim of interest to you personally
- Perform a regression analysis using R Studio
- Clearly and efficiently present your findings through writing a semi-formal statistical report, an oral discussion, and academic poster presentation
- Implement a strategy for learning an unfamiliar regression topic
- Establish and maintain effective working relationships with others
- Develop an understanding of the obligations of a statistician for reporting findings ethically

What tools do you need to be successful?

- **Canvas:** All of the resources you should need or want for this course will be posted to Canvas. This site can also be used to communicate with your peers, upload assignments, and download supplemental notes. Additionally, announcements will be posted via Canvas. It is your responsibility to check for course updates through email or announcements.
- **Textbook**: The primary recommended text for this course is *A Second Course in Statistics: Regression Analysis (7th ed) ISBN:* 978-0321831453. There is a more recent

edition, but the 7th will suffice. The secondary recommended text for this course is *Practicing Statistics: Guided Investigation for the Second Course ISBN: 978-0321586018***.**
    - The sections listed in the syllabus correspond with the *Regression Analysis* book.
    - While I will provide supplemental notes, I recommend you read the textbook to create your own collection of notes.
- **R Studio:** The software introduced and featured in this course is R. RStudio is a software that runs R with additional user-friendly features. RStudio will be used for all in-class R demonstrations and examples. Both R and RStudio are free for download for all operating systems. R will need to be downloaded first. RStudio Desktop is the recommended version for download. Be sure to download the version that match the specifications of your computer. **There was an update to Base R in the fall, you will want to download that new version (4.5.2 "[Not] Part in a Rumble")**
- **Gradescope:** You will use Gradescope to submit most assignments. Please familiarize yourself with the appropriate submission procedure. If an assignment is not submitted correctly, it will not be graded.

Who/What is here to help you be successful?

- **Instructor**: Krista Varanyak (that's me). My research is Statistics Education (how to teach statistics) and I am currently learning more about intermediate stats courses across the country. When I am not teaching (,or prepping, or grading, or meeting with students,…) I enjoy running, reading, and remodeling my home.
    - **Email**: kristav@virginia.edu (same as kmv9q@virginia.edu)
    - **Office**: Halsey 207
- **Undergraduate Course Assistants**: Great resources in each class meeting + office hours
    - 8:00am Section 001 – Laine Giovanniello (bhn7bm@virginia.edu)
    - 9:30am Section 002 – Mikayla Carter (jvc8dg@virginia.edu)
    - 11:00pm Section 003 – Cameron Barlow (kvr5dj@virginia.edu)
- **Office Hours:** You may attend any of the hours listed below or email Prof V to schedule an appointment. Please include your availability when sending an email.
    - **TBD**
- **YOU (Discussion Boards):** I think students are sometimes the best source of information, so we will utilize Discussion Boards as our question forum to communicate outside of class. You can post and reply to questions about content, assignments (except where noted), R Studio support, and more. The course assistants and I will monitor these about once a day as well.
- **Internet Sources**: Google searching is a skill and will be helpful for troubleshooting and learning new skills. We cover some of the basics, but you may also want to refer to some R Studio handy sheets (Base R, R Markdown, Graphics, and many others)

How will you achieve the goals?

Ultimately, developing the skills required to think and act like a statistician comes down to **learning** the course material, **recalling** the course material, and **applying** the course material. Therefore, in order to demonstrate you are meeting the course objectives, which lead to the overall course goal, you will be assessed on your demonstration of content engagement, your content knowledge, and content application.

- **Content Engagement:** The purpose of each aspect of content engagement is to demonstrate that you are learning the course material. Activities are designed for you to engage with the material to help you discover underlying themes in regression. The primary purpose of these activities is to inform me about your learning and give an opportunity for you to reflect on your own progress.
    - Pre-Class Readings & Videos: Preparing for each class meeting (reading the relevant textbook sections, watching lecture videos, etc.) will help you get the most out of the class meeting. You are expected to read the appropriate sections listed on the calendar and be prepared to answer the unit questions and watch the pre-recorded lecture video (usually about 30 minutes). **Each video will have embedded questions that will be marked towards your grade and are due prior to the class meeting which the content is covered.**
    - Live Class Meeting: Each class meeting will be spent reviewing an example from the corresponding pre-recorded lecture. There will also be time to answer questions from the lecture and work in groups. Class meetings are not recorded. **Class attendance and engagement will be taken each meeting through various polling metrics and scored towards your grade.**
        - Attendance: Attendance is recorded in this class. You are permitted 2 absences without penalty. Please refer to the attendance section at the end of the syllabus for the details for additional information.
        - Assignments: The goal of these activities is to expose you to the nuances of the regression process, develop your own understanding of the techniques, and provide low-stakes situations for you to explore techniques through practice. These activities will vary in structure for each class meeting (individually or in groups; may use R; may be discussion or written; etc). Most of these assignments will be completed in the class meeting, but there may be a few you will complete outside of class.
    - Practice Set: There will be 4 unit practice sets. These will not be graded, nor collected. The purpose it to provide you additional exercises to help prepare for exams. Solutions will be posted.
- **Content Knowledge**: Content knowledge assessments are designed for you to show me what you know through exams. The purpose of each content knowledge assessment is to demonstrate that you can recall the course material and perform related tasks correctly.
    - Exams: There will 2 exams, each covering 2 units of material. The purpose of the exams is for you to show you can correctly apply the techniques discussed in

units. Further, you will demonstrate your ability to apply and extend the concepts we discuss through an exploration and analysis of a real-world problem.

- Exams will consist of multiple choice and free response questions and will be completed during your class meeting. Exams cannot be made up or retaken for any reason.
- Exam 1- Tuesday, February 17th
- Exam 2- Thursday, March 26th

- **Content Application**: The purpose the of content application is to demonstrate that you can apply and synthesize the course material. The project is designed to expose you to the real-world intricacies of being a statistician through application. As a group of 3-4 students you will select a research topic (and data set) of interest, perform a statistical analysis, then communicate your findings through a written report, and/or a presentation. **(Specific details to follow).**
  - Final Project: You will use methods covered in the exploratory data analysis phase to collect (or compile) your own data and present it in a thoughtful way without inference. Then, you will apply and extend the multiple linear regression and logistic regression techniques covered and present in an academic poster session and report.

How will your objectives be measured?

- Pre-lecture Video Quizzes + Material (10%): These are questions imbedded into the pre-class videos that you will find on the canvas modules. They must be completed by the start of the corresponding class meeting. Your grades will be averaged out over the term where 90-100 = 10 points; 85-89.9= 9 points; 80-84.9 = 8 points; 75-79.9=7, 70-74.9=6 points; 60-69.9 = 5 points; 50-59.9=2points; 0-50=0points
- Assignments (20%): This score will be comprised of: your class assignments, scores on various project workday assignments, and out of class assignments. If you miss class the day something is due, you can still submit the assignment for full credit by the deadline. If you are unable to submit the activity on time, you cannot make it up for credit, but it should still be completed for your own understanding. To account for short-term illnesses, emergencies, interviews, etc, classwork scores will be weighted to allow for 2 class meetings throughout the term to be missed without penalty.
    - Your classwork is scored out of points, not per assignment. Therefore, your lowest scores are not excluded from your grade, instead your score will be measured out of 200 points less than what was submitted. Assignments range from 50-200 points, and there is typically a total of 900-1200 points on assignments throughout the semester.
        - For example, if 1200 points of work is collected and you earned 1100 points, your score will be 1100/1000 = 100% (capped scoring)
- Exams (40%): Each exam will be graded with an emphasis on correct logic in addition to a correct final solution. Exams cannot be made up, please mark these dates on your calendar. Your higher exam score is scored 30% and your lower exam score is weighted 10%.
- Project (30%): Details and rubrics will be given at the date listed on the calendar.
- Attendance: Attendance is required in this course. Beginning 1/20, you are permitted 2 automatically-excused absences, no documentation is needed to receive this. If you miss beyond the allotted permitted absences, your course grade will lower by 1% per additional absence.
    - Please see the attendance policy section for further details on particular university- approved accommodations and extenuating circumstances.
- Grade Thresholds: The default university thresholds are used in this course, for more information see https://virginia.service-now.com/its?id=itsweb_kb_article&sys_id=1153c16fdba41f444f32fb671d961934#undergraduate

When will we do all this? – need to update chapters to correspond with other book.

Day	Date	Material Covered (Pre Class Video)	Chapter/Topic (Class Activity)	Work to Submit
Unit 1: WHAT IS A MODEL? (Ch 2-4.11)				
Tuesday	13-Jan	Welcome, Syllabus, Policies,	**Course Introduction**	
Thursday	15-Jan	Unit 1.1: Ch 2-3; Introduction to R Studio	**What is model?** Simple linear regression model	
Tuesday	20-Jan	Unit 1.2: Ch 4.1-4.5 +4.11	**What is a model not?** Estimating the parameters (betas and sigma) of regression.	
Thursday	22-Jan	Unit 1.3 Ch 4.6-4.8	**What makes a good model?** Assessing the model; parameter inference, R^2 vs adjusted R^2, Global F	
Tuesday	27-Jan	Unit 1.4: Ch 4.9	**How do you use a model?** Prediction vs Estimation	
Unit 2: HOW TO… BUILD A MODEL? (Ch 4.10-5.10)				
Thursday	29-Jan	Unit 2.1: Ch 4.12, 5.7-5.9	**How to… turn colors into numbers?** Qualitative variables	
Tuesday	3-Feb	Unit 2.2: CH 4.10-4.13	**How to… change the slopes?** Interactions	
Thursday	5-Feb	Unit 2.3: CH 5.11	**How to… Build a model?** Nested Test, Model Validation	

Tuesday	10-Feb	Unit 2.4: Ch 7.4 + Ch 6.1-6.2	**What happens when "eXes" aren't independent?** Multicollinearity + Stepwise Regression	
Thursday	12-Feb		**Unit 1-2 Wrap**	
Tuesday	17-Feb		**Exam 1: Unit 1-2 Exam**	
Unit 3: WHY ARE ERRORS "NORMAL"? (Ch 8 + 7.6)				
Thursday	19-Feb	Unit 3.1: Ch 8.1-8.5, 8.7	**Why are errors "Normal"?** Regression Residuals & assumptions	
Tuesday	24-Feb	Unit 3.2: Ch 8.6	**Can one observation change a model?** Outliers and Influential Points	
Thursday	26-Feb	Unit 3.3: Ch 7 + P-values	**Friends don't let friends extrapolate.** Regression Pitfalls	
Tuesday	3-Mar	Spring Break	**NO CLASS**	
Thursday	5-Mar	Spring Break	**NO CLASS**	
Unit 4: WHAT ARE THE ODDS? (CH 9.5-9.6, CH 12, Supplemental)				
Tuesday	10-Mar	Unit 4.1: Ch 9.5-9.6, Supplemental	**Why not use linear regression?** Logistic Regression	
Thursday	12-Mar	Unit 4.2: Supplemental	**What are the odds?** Logistic regression cont	
Tuesday	17-Mar	Unit 4.3: Ch 12.1-12.3, 12.5	**One factor. Many means. One test.** One factor ANOVA	
Thursday	19-Mar	Unit 4.4: Ch 12.7 & 12.9	**ANOVA Follow Up** ANOVA assumptions and post hoc analysis	

Tuesday	24-Mar		**Unit 3-4 Wrap**	
Thursday	26-Mar		**Exam 2: Unit 3-4 Exam**	
Final Project				
Tuesday	31-Mar	Project Assigned+ data librarians	**Project Work Day**	
Thursday	2-Apr		**Project Work Day**	
Tuesday	7-Apr		**Project Work Day**	
Thursday	9-Apr		**Project Work Day**	
Tuesday	14-Apr		**Project Work Day**	
Thursday	16-Apr		**Project Work Day**	
Tuesday	21-Apr		**Project Work Day**	
Thursday	23-Apr		**Project Presentations**	
Tuesday	28-Apr		**Project Presentations**	

There is no final exam for the course, and we will not use our exam period.

All work will be completed and submitted by the last day of the semester.

How will we all –y'all and I- work together?

In order to create a positive, equitable learning environment for all students, the following policies are in place and resources are available:

- **Classroom etiquette**-There are no right or wrong answers in Statistics (most times). You are free to ask questions related to our topic and are to treat your peers and instructor with respect.
- **Technology Policy**- During the lecture portions of this course, **we will promote a distraction free, engaging course environment through "laptop free" learning**. You are permitted to use tablets for notetaking or accessing course material only, as those technologies do not cause as rigorous distractions.
    - It is suggested you print any associated note prior to class and keep track in a binder or course folder. I will inform you if I will provide any specific handouts or guided notes.
    - You will be permitted to use your laptops during group work sessions as you will need to code within R Studio.
    - Phones should be turned on do not disturb. You may need your phone to submit to a class poll.
    - Headphones should be off while you are in class for the entire session.
    - If you have an accommodation that goes against these course policies, please meet with me so we can ensure you will be set up for success.
- **Honor Policy**- By enrolling in this course, you agree to uphold the Honor Policy of the University. Policies are laid out in this syllabus for each type of assignment. Additionally, you are not to share any assignments, or materials including, but not limited to classwork, quizzes, homework, or projects with students from previous or future semesters. You are not permitted to use online question forums as listed in this syllabus. If you have any questions about the policies, it is always better to ask than assume.
    - **Generative AI**
      Students may use generative AI programs, including ChatGPT, on assignments with some limitations. Generative AI programs are not a replacement for human creativity, originality, critical thinking, and research. Writing, whether text or code, is a craft that must develop over time to refine the author's individual voice, though, with proper attribution, AI programs may be used as a tool in this process.
    - **Students are NOT permitted to submit any information provided by the instructor, including whole or partial problems, prompts, background information, examples, data, etc., into generative AI programs.**
    - Any submitted work using AI programs must clearly indicate what part is the student's work and what part is generated by the AI programs with appropriate notation and citations. In such cases, no more than 25% of the submitted work should be generated by AI programs. Students should also indicate how generative AI programs informed their process and their submitted work, including how the generated information was validated.

- Assignment instructions will provide additional guidance as to how students should provide transparency about the use of generative AI programs in submitted work. Using a generative AI program without proper attribution is a violation of the Honor Code and will be treated as such.
- Students using generative AI programs should view them as a collaborative tool, understanding that they can be trained on limited datasets that may be out of date. Additionally, generative AI datasets are trained on pre-existing material, that may include copyrighted material and, therefore, relying on generative AI programs may result in plagiarism or copyright violations. Finally, students should keep in mind that the goal of generative AI programs is to produce content that seems to have been produced by a human, not to produce accurate or reliable content and, therefore, relying on generative AI programs may result in submission of inaccurate content. It is the student's responsibility – not the program's – to assure the quality, integrity, and accuracy of submitted work.

- **Make Ups**- The general policy for all assignments and class meetings is that they are not to be made up and will not be accepted after the due date. Thus if for any reason you will miss class (including short-term illness, overloaded weeks, interviews, or other temporary unforeseen circumstances.) If you do not submit an assignment on time to Gradescope (or whatever platform) please do not send it to me or the grader through email- it will not be graded.
  - In the first few weeks of class, you should find at least two people who you can check in with in the event that you are ever ill. I strongly recommend that at least one of these people is NOT in your classwork group.
  - If you know you will miss an extended period of class due to illness or another reason, please reach out and we will make arrangements to provide a reasonable accommodation for you as you recover. For these situations, you should also be in contact with your Advising Dean, Academic Advisor, and/or and SDAC Advisor. Case-by-case exceptions will not be granted beyond these university channels.
- **Attendance-** Attendance is recorded each day via Poll Everywhere, which must be connected to your UVA Netbadge Account. Completing the poll while not present in the classroom is grounds for academic dishonesty and will be treated as such.
  - Errors in attendance record can occur, though infrequently. You have one week from the class period to raise a potential attendance error with the professor. Otherwise it will be recorded as an absence.
  - Beginning 1/20, Each student is permitted 2 absences. These do not require documentation and cover any reason you may miss class (sick, interview, vacation, slept in, car break down, etc). Absences that fall in this category, beyond 2 will result in a 1% lowering of your course grade per absence.
  - If an extenuating circumstance arises (long term hospitalization), you should contact the professor ASAP to discuss if your absences may be excused.

    - If you know ahead of time you will miss a particular class for a university sanctioned reason or religious holiday, please meet with your professor ASAP to make appropriate arrangements.
    - If you find that you will miss more than 2 weeks of the semester, you should schedule a meeting with your professor to determine whether an action plan is possible to put into place to preserve the integrity of the course's learning goals. This may also involve a conversation with your advising dean.
- **Timely Arrival-** Our class meetings begin at 8:00am, 9:30am, 11:00am respectively. You are expected to attend only your enrolled section and to remain in class the entire 75 minute duration. You should be in class, and ready to begin by the start of the class meeting. Arriving regularly well after our start time is disruptive and disrespectful to myself and your peers.
    - If you arrive more than 15 minutes late or leave more than 15 minutes early, you will be marked as absent.
    - If you regularly arrive late, or leave early (but within 15 minutes), this may count towards your absences at my discretion.
        - You should be courteous and respectful. I understand occasionally, needing to arrive a few minutes late, but making this a regular occurrence is not advisable nor professional.
        - If you plan to leave class early - It would be appreciated to let me know prior to the start of class. Otherwise, I assume there is a concern.
        - You should also be aware that if you miss any portion of the class, you might miss pertinent course information.
    - Read the university policy on attendance: https://college.as.virginia.edu/class-attendance
        - To elaborate on remaining engaged in class: you can do so by answering questions, participating in discussion with your group, and not working on other courses or outside materials while you are in class. Your participation is my primary way to engage your real-time understanding to properly pace the class.
- **Office hours policies:** The following policies will apply to all office hours to be able to answer as many questions for as many students as possible.
    - All office hours will be open-door with a queue. To discuss personal situations, students should schedule an appointment with the instructor.
    - Students should arrive prepared with specific questions.
    - Students should not expect that variations of the questions “Is this correct?”' or “Will this earn full credit?” will be answered.
    - Students should not expect that significant portions of material from a missed class will be explained.
- **Email**- Before sending me or a CA an email, you are expected to consult with at least three other sources to find your answer (Peer, Syllabus, textbook, R help pages, Discussions). Most questions should be posted and answered within Discussion Boards. Assuredly, you are not the only one with that question. If you have a personal question

or would like to set up a meeting, please see Canvas for the best way to communicate with me. Emails that are answered in the syllabus, about a specific assignment grading, or submitting a late assignment without proper accommodations may not receive a reply. Emails beyond these will receive a reply within 2 business days.

- **SDAC**- All students with special needs requiring accommodations should present the appropriate paperwork from the Student Disability Access Center (SDAC). It is the student's responsibility to present this paperwork in a timely fashion and to follow up about the accommodations being offered.

Note: The syllabus is subject to change. All changes will be submitted in writing.