STA 210: Regression Analysis

Learn approaches for analyzing multivariate data sets, emphasizing analysis of variance, linear regression, and logistic regression. Learn techniques for checking the appropriateness of proposed models, such as residual analyses and case influence diagnostics, and techniques for selecting models. Gain experience dealing with the challenges that arise in practice through assignments that utilize real-world data. This class emphasizes data analysis over mathematical theory.

Course info

Lectures

🏛 Gross Hall 103   📅 Mon and Wed 10:05a - 11:20a

Labs

Lab 01   🏛 Link Classroom #5   📅 Thu 3:05p - 4:20p

Lab 02   🏛 Link Classroom #5   📅 Thu 4:40p - 5:55p

Lab 03   🏛 Link Classroom #1   📅 Thu 4:40p - 5:55p

Teaching team and office hours

Instructor	Prof. Maria Tackett (http://stat.duke.edu/~mt324/)	✉ (mailto:maria.tackett@duke.edu) ⬡ (https://github.com/matackett)	Wed 3p - 5p	Old Chem 118B
TAs	Youngsoo Baek (https://stat.duke.edu/people/youngsoo-baek-0)	✉ (mailto:youngsoo.baek@duke.edu) ⬡ (https://github.com/ybaek)	Mon 1p - 3p	Old Chem 203B
	Cody Coombs (http://linkedin.com/in/cody-coombs-3b8034158)	✉ (mailto:cody.coombs@duke.edu) ⬡ (https://github.com/coombscody)	Tue 1p - 3p	Old Chem 203B
	Sophie Dalldorf (https://www.linkedin.com/in/sophie-dalldorf-598a16192/)	✉ (mailto:sophia.dalldorf@duke.edu) ⬡ (https://github.com/sophiedalldorf)	Fri 1p - 3p	Old Chem 203B
	Jonathan Klus (https://stat.duke.edu/people/jonathan-klus)	✉ (mailto:jonathan.klus@duke.edu) ⬡ (https://github.com/jonklus)	Mon 3p - 5p	Old Chem 203B
	Matty Pahren (https://www.linkedin.com/in/mattypahren)	✉ (mailto:martha.pahren@duke.edu) ⬡ (https://github.com/mpahren)	Tue 3p - 5p	Old Chem 203B
	Ethan Shen (https://www.linkedin.com/in/ethan-shen-931010134/)	✉ (mailto:ethan.shen@duke.edu) ⬡ (https://github.com/ethann-shen)	Wed 5p - 7p	Old Chem 203B

Textbooks

Handbook of Regression Analsyis (http://sakai.duke.edu)	James, Witten, Hastie, Tibshirani	Springer, 1st edition, 2013
R for Data Science (http://r4ds.had.co.nz/)	Chatterjee, Simonoff	Wiley, 1st edition, 2013

DUKE STATSCI (HTTP://STAT.DUKE.EDU/)   RSTUDIO (HTTPS://RSTUDIO.CLOUD/)

GITHUB (HTTPS://GITHUB.COM/STA210-SP20)   (./)

Syllabus

Click here (syllabus.pdf) to download a PDF copy of the syllabus.

Course Learning Objectives

By the end of the semester, you will...

- know methods for analyzing multivariate datasets and understand how these methods are applied in research
- know how to check whether proposed statistical models are appropriate for given data
- develop proficiency in addressing complex research questions using statistical evidence
- develop proficiency in computing tools used to conduct reproducible statistical analyses, specifically R and Git
- understand the process of data-based research by working on an independent research project

Activities & Assessments

The following activities and assessments will help you successfully achieve the course learning objectives. By experiencing the course content in different ways, you will not only gain a better understanding of regression analysis, but you will also get experiences that can guide you as you apply what you've learned in future academic and professional projects.

Homework (25%)

In homework, you will apply what you've learned during lecture and lab to complete data analysis tasks. You may discuss homework assignments with other students; however, homework should be completed and submitted individually. Homework must be typed up using R Markdown and submitted in the appropriate GitHub repository.

Individual homework extensions will only be given for extenuating circumstances. There must be notification from your academic dean about the need for an extension before the homework deadline. We do not grant extensions after the assignment is due.

To accommodate unexpected and temporary situations (e.g. interviews and other travel, unusually busy weeks, short term illness, etc.), the lowest homework grade will be dropped at the end of the semester.

Labs (15%)

In labs, you will apply the concepts discussed in lecture to various data analysis scenarios, with a focus on the computation. You will work on lab assignments in teams, and all team members are expected to contribute equally to the completion of each assignment. You are expected to use the team's repository on the course's GitHub page as the central platform for collaboration. Commits to this repository will be used as a metric of each team member's relative contribution for each lab. You will also be asked to evaluate your team members' performance periodically during the semester.

A portion of each lab grade will be for attending and participating during the lab session. You must attend lab and participate during the lab session to receive credit for the group's lab assignment. If you did not attend lab or did not participate during the lab session, your name should not be included in the final group submission.

If you miss lab, you should complete lab for partial credit by the regular lab deadline.

Individual lab extensions will only be given for extenuating circumstances. There must be notification from your academic dean about the need for an extension before the lab deadline. We do not grant extensions after the assignment is due.

To accommodate unexpected and temporary situations (e.g. interviews and other travel, unusually busy weeks, short term illness, etc.), the lowest lab grade will be dropped at the end of the semester.

Exams (Exam 01: 20%, Exam 02: 20%)

The exams are an opportunity to assess the knowledge and skills you've learned. They will include both the conceptual and mathematical and conceptual aspects of regression. Both exams will be given during a lecture class period, and you will be permitted to bring one page of hand-written notes to each exam. The exam dates are as follows:

- Exam 01: Wednesday, February 26

- Exam 02: Wednesday, April 15

Exam dates cannot be changed and no make-up exams will be given. If you must miss an exam, your absence must be officially excused before the exam due date. If your absence is excused, the missing exam grade will be imputed at the end of the semester based on your performance on other relevant individual coursework.

Final Project (15%)

The purpose of the project is to apply what you've learned throughout the semester to analyze an interesting data-based research question using regression. The project will be completed in teams, and each team will present their results during the final exam period, Tuesday, April 28, 9a - 12p. **Project presentations must be done during the specified period.**

Teamwork & Engagement (5%)

This part of the final grade will consist of three components.

- Team Feedback Surveys. You will be assigned to diverse teams of 3-4 students based on the results of a Get to Know You survey at the beginning of the semester. These teams will stay consistent throughout the semester (barring extraordinary circumstances). You will primarily work in these teams on lab assignments, the final project, and some in-class activities.

- Daily Engagement Surveys. After each class you will receive a quick survey of consisting of three questions about engagement in the day's lecture. You are expected to fill out the survey even if you did not attend lecture that day. You are required to complete at least 80% of the daily engagement surveys along with the pre-course and post-course surveys to receive full credit for this component of the grade.

- In-Class Questions. We will use active learning in class, which will include answering questions through a web browser. You are required to complete at least 75% of these in-class questions to receive full credit for this component of the grade. You will not be graded based on accuracy; however, you should make a earnest attempt at answering the questions, since they will be used in part to give real-time feedback about the class's understanding of the lecture material. You will be identified using your Net Id. It is your responsibility to ensure your Net Id is entered correctly on each survey to receive credit. In-class questions cannot be made up; there is built-in flexibility so you can earn full credit even if you miss multiple lectures.

Grade Calculation

The final grade will be calculated as follows:

Homework	25%
Labs	15%

Exam 01	20%
Exam 02	20%
Final Project	15%
Teamwork & Engagement	5%

Class attendance in lecture and lab is a firm expectation; frequent absences or tardiness will be considered a legitimate cause for grade reduction.

If you have a cumulative numerical average of 90 - 100, you are guaranteed at least an A-, 80 - 89 at least a B-, and 70 - 79 at least a C-. The exact ranges for letter grades will be determined at the end of the semester.

Diversity & Inclusion

It is my intent that students from all diverse backgrounds and perspectives be well-served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that the students bring to this class be viewed as a resource, strength and benefit. It is my intent to present materials and activities that are respectful of diversity and in alignment with Duke's Commitment to Diversity and Inclusion (https://provost.duke.edu/initiatives/commitment-to-diversity-and-inclusion). Your suggestions are encouraged and appreciated. Please let me know ways to improve the effectiveness of the course for you personally, or for other students or student groups.

Furthermore, I would like to create a learning environment for my students that supports a diversity of thoughts, perspectives and experiences, and honors your identities. To help accomplish this:

- If you feel like your performance in the class is being impacted by your experiences outside of class, please don't hesitate to come and talk with me. If you prefer to speak with someone outside of the course, your academic dean is an excellent resource.
- I (like many people) am still in the process of learning about diverse perspectives and identities. If something was said in class (by anyone) that made you feel uncomfortable, please talk to me about it.

Accessibility

Duke University is committed to providing equal access to students with documented disabilities. Students with disabilities may contact the Student Disability Access Office (SDAO) (https://access.duke.edu) to ensure your access to this course and to the program. There you can engage in a confidential conversation about the process for requesting reasonable accommodations both in the classroom and in clinical settings. Students are encouraged to register with the SDAO as soon as they begin the program. Please note that accommodations are not provided retroactively. More information can be found online at access.duke.edu (https://access.duke.edu) or by contacting SDAO at 919-668-1267, SDAO@duke.edu (mailto:sdao@duke.edu).

This class will use the Testing Center to provide testing accommodations to students registered with and approved by the SDAO. The center operates by appointment only and appointments must be made at least 7 consecutive days in advance, but please schedule your appointments as far in advance as possible. You will not be able to make an appointment until you have submitted a Semester Request with the SDAO and it has been approved. For instructions on how to register with SDAO, visit their website at https://access.duke.edu/requests (https://access.duke.edu/requests). For instructions on how to make an appointment at the Testing Center, visit their website at https://testingcenter.duke.edu (https://testingcenter.duke.edu).

Getting Help

- If you have a question during lecture or lab, feel free to ask it! There are likely other students with the same question, so by asking you will create a learning opportunity for everyone.
- The teaching team is here to help you be successful in the course. You are encouraged to attend office hours during the times posted on the home page to ask questions about the course content and assignments. A lot of questions are most effectively answered in-person, so office hours are a valuable resource. Please use them!
- Outside of class and office hours, any general questions about course content or assignments should be posted on Piazza. There is a chance another student has already asked a similar question, so please check the other posts on Piazza before adding a new question. If you know the answer to a question posted on Piazza, I encourage you to respond!

There are times may need help with the class that is beyond what can be provided by the teaching team. In those instances, I encourage you to visit the Academic Resource Center. The Academic Resource Center (ARC) (https://arc.duke.edu) offers free services to all students during their undergraduate careers at Duke. Services include Learning Consultations, Peer Tutoring and Study Groups, ADHD/LD Coaching, Outreach Workshops, and more. Because learning is a process unique to every individual, they work with each student to discover and develop their own academic strategy for success at Duke. Contact the ARC to schedule an appointment. Undergraduates in any year, studying any discipline can benefit! Contact ARC@duke.edu (mailto:arc@duke.edu), 919-684-5917, 211 Academic Advising Center Building, East Campus – behind Marketplace.

Course Materials

All lecture notes, assignment instructions, up-to-date schedule, and other course materials may be found on the course website, http://bit.ly/sta210-sp20 (http://bit.ly/sta210-sp20). I will periodically send announcements via email, so please check your email regularly. You may also find a copy of announcements on Sakai.

Late Work

Homework or lab assignments submitted late but within 24 hours of the deadline may be accepted with a 20% penalty. Homework or lab assignments submitted any later will <u>not</u> be accepted.

Late work will not be accepted for the exams or the final project.

Academic Honesty

By enrolling in this course, you have agreed to abide by and uphold the provisions of the Duke Community Standard (https://studentaffairs.duke.edu/conduct/about-us/duke-community-standard) as well as the policies specific to this course. Any violations will automatically result in a grade of 0 on the assignment and will be reported to Office of Student Conduct (https://studentaffairs.duke.edu/conduct) for further action.

- You may not discuss or otherwise work with others on the exams. Unauthorized collaboration or using unauthorized materials will be considered a violation for all students involved. More details will be given closer to the exam date.
- **Reusing code**: Unless explicitly stated otherwise, you may make use of online resources (e.g. StackOverflow) for coding examples on assignments. If you directly use code from an outside source (or use it as inspiration), you must explicitly cite where you obtained the code. Any recycled code that is discovered and is not explicitly cited will be treated as plagiarism.
- On individual assignments, you may not directly share code or write up with other students. On team assignments, you may not directly share code or write up with another team. Unauthorized sharing of the code or write up will be considered a violation for all students involved.

Regrade Requests

Regrade requests should be submitted through the regrade request from on Gradescope. **Requests for a regrade must be made within a week of when the assignment is returned; requests submitted later will not be considered.** You should only submit a regrade request if there is an error in the grade calculation or a correct answer was mistakenly marked as incorrect. You should not submit a regrade to dispute the number of points deducted for an incorrect response. *Please note that by submitting a regrade request, your entire assignment may be regraded and you may potentially lose points.*

Due to the time consuming nature of responding to regrade requests, you should attend office hours and ask a member of the teaching team about the feedback before submitting the request When you submit a request, please indicate which member of the teaching team you spoke with. *Note: Grades can only be changed by the Professor Tackett. Teaching Assistants cannot change grades on returned assignments.*

No grades will be changed after the final project presentations.

Technology

Cell phones and other electronic devices should be turned off or put on silent during class. When you use a laptop, tablet, or comparable device, please ensure that the volume set to mute and the device is only used for class purposes. In general, you should focus on the class discussion/activity at hand and refrain from engaging in other work or outside activities.

DUKE STATSCI (HTTP://STAT.DUKE.EDU/)   RSTUDIO (HTTPS://RSTUDIO.CLOUD/)

GITHUB (HTTPS://GITHUB.COM/STA210-SP20)   (./)