

SDS358: Applied Regression Analysis
Department of Statistics and Data Sciences – Fall 2018
Unique: 56835
PHR 2.116, MWF 10:00am – 11:00am

“You can’t go from knowing nothing to becoming an expert without going
through a period of great frustration & great suckiness.”
– Hadley Wickham (creator of ggplot)

Instructor	Dr. Michael J. Mahometa
Email	michael.mahometa@austin.utexas.edu
Office	GDC 7.406
Office Hours	Scheduled: Tues 9:00am-11:00am OR by appointment (confirmed by email)
Course Website	https://canvas.utexas.edu

COURSE DESCRIPTION

This is intended to be a hands-on course. Through software application, discussion, and guided instruction, students will first learn correlation and simple linear regression – what data is appropriate, how to run the analysis, and how to interpret the output. Then, students will move on to multiple regression with combinations of predictor variables, both continuous and categorical. There will be a discussion/application of ANOVA, prior to proceeding on to logistic regression: the prediction of discrete events.

Because of the hands-on nature of the course, there will be opportunity to apply the skills learned in class to real data for a final project. Students will learn how to use software (R and RStudio) in a weekly lab environment, to integrate course knowledge. There will be weekly assessments, 3 major exams, and a final poster project to be presented at the end of the semester.

COURSE OBJECTIVES

Lectures and Labs will provide students with a working knowledge of regression, and at the completion of the course be able to:

- 1) Determine if a research question can be answered using some form of regression
- 2) The steps (and pitfalls) of running the regression analysis to answer the research question
- 3) Interpret the output of the analysis in a concise manner, capable of being used in publication.

COURSE STRUCTURE

Each week, we will cover a basic topic of regression analysis, from correlation to moderation to logistic regression. Each topic will build on the prior topics, so understanding is crucial. Each week will be divided into Class and Lab (Class on MW, Lab on F), with an independent assignment following each Lab. We will have UNIT assessments that will test your knowledge through several weeks, and a final project that will cover all your skills.

PREREQUISITES

Knowledge of algebra, completion of an Introductory Statistics Course

REQUIRED TEXT AND MATERIALS

Textbook: NOT REQUIRED

Stat2: Building Models for a World of Data, Cannon, A.R. et al.

Available for purchase at the UT Bookstore Co-Op

On reserve (two hour periods) at the PCL

Software:

R and RStudio

R: <https://www.r-project.org>

RStudio: <https://www.rstudio.com/products/rstudio/download/>

SDSRegressionR package available on GitHub (instructions on Canvas).

Calculator with a square root function

COURSE EXPECTATIONS

You are responsible for ensuring that you are prepared for both Class and Lab. Here are the daily requirements for Class and Lab:

1. **Arrive on Time.** Please be courteous to your peers. I understand that the 40 acres is big—but please do your best.
2. **Put Away Electronic Devices.** Put your cell phone in silent mode and put away any other electronic devices (e.g., computers, iPods, iPads, etc.). You'll only need the computers on Lab days. You **will** need access to a machine with R and RStudio for your weekly assignments and Labs.
3. **Bring Your Calculator.** You are expected to bring a scientific or graphing **calculator** to lecture each day.
4. **Participate:** You will be called on at various points in the Class and Lab. You will be graded on how well you participate. You need not have the right answer, just be willing to engage. Working through a problem out loud is completely acceptable. Participation *is part of your grade*.
5. **Respect:** Class is a safe place to engage and share ideas. I expect that you will respect not only me, but also your classmates, and yourself.

ASSESSMENTS

Weekly Assessments:

Each week, students will need to complete the following:

- A Canvas reading quiz to assess knowledge prior to instruction
- Two lecture periods to cover main topics
- One lab period to apply topics to a dataset, answering a research question
- A take home assessment (done independently) that will test both Lecture and Lab knowledge.

Unit Assessments:

There will be three UNIT assessments (see COURSE SCHEDULE) that will test both your theoretical knowledge and application knowledge. All or a portion of these assessments will be completed in Canvas.

Final Project:

You will complete a final project, to be presented at the class's designated Final Time. This will be a poster presentation of an advanced regression analysis (you must clear your topic, technique, and data with the instructor by Week 10) answering a research question of your choice. You will also be responsible for submitting a Final Project Proposal, outlining your research question, your data source, and your suggested technique. Project ideas cannot be changed once the Final Project Proposal is submitted. You may collect data from any place you see fit. Possible data sources will be posted on Canvas as the semester progresses. The Final Project is mandatory for the course.

GRADING

Grading of assessments are as follows:

Assessment	Points	Frequency	Total Semester Points
Participation	100	1	100
Reading Quizzes	10	13	130
Weekly Labs	50	12	600
Unit Tests	100	3	300
Final Project Proposal	50	1	50
Final Project	220	1	220
TOTAL			1400

FINAL COURSE GRADE

Final grades will be assigned according to the grade cutoffs listed). Grades are based on point values, and cutoffs are firm. There is no curve or extra credit in this course. Keep track of your progress.

Assigned Course Grade	Points Required
A	1302
A-	1260
B+	1218
B	1162
B-	1120
C+	1078
C	1022
C-	980
D+	938
D	882
D-	840

CLASS POLICIES

Course Materials

ALL course materials provide to the student is copyright of their respective owner. No material that you receive from this course may be reproduced in any form.

Absences

Absences due to illness must be supported by a doctor's note for the time frame effected. All other Absences, including those academically related, must be shared with the professor at least two-weeks prior to the effected class dates. Failure to do so will result in loss of participation points, or related assignment points.

Religious Holy Days

By UT Austin policy, you must notify me of your pending absence at least fourteen days prior to the date of observance of a religious holy day. If you must miss a class, an examination, a work assignment, or a project in order to observe a religious holy day, I will give you an opportunity to complete the missed work within a reasonable time after the absence.

Q Drop Policy

If you want to drop a class after the 12th class day, you'll need to execute a Q drop before the Q-drop deadline, which typically occurs near the middle of the semester. Under Texas law, you are only allowed six Q drops while you are in college at any public Texas institution. For more information, see: <http://www.utexas.edu/ugs/csacc/academic/adddrop/qdrop>

Student Accommodations

Students with a documented disability may request appropriate academic accommodations from the Division of Diversity and Community Engagement, Services for Students with Disabilities, 512-471-6259 (voice) or 1-866-329-3986 (video phone). <http://ddce.utexas.edu/disability/about/>

- Please request a meeting as soon as possible to discuss any accommodations
- Please notify me as soon as possible if the material being presented in class is not accessible
- Please notify me if any of the physical space is difficult for you

Academic Integrity

Each student in the course is expected to abide by the University of Texas Honor Code:

“As a student of The University of Texas at Austin, I shall abide by the core values of the University and uphold academic integrity.”

You will be asked to do assignments out of class. Your work should be your work – period. This extends to the idea of “helping” one another. Helping on a graded assignment is cheating. The only person who is permitted to help you on a graded assignment is the instructor. If it is determined that any portion of an assignment was completed in a non-independent manner, all parties will receive zeros for the entire assignment.

You are responsible for understanding UT's Academic Honesty Policy which can be found at the following web address: http://deanofstudents.utexas.edu/sjs/acint_student.php

LETTERS OF RECOMMENDATION

Some of you may wish to ask me for a Letter of Recommendation, to use for graduate school or employment applications. I take this very seriously. But, you need to know that professors are not obligated to provide a Letter of Recommendation just because a student asks. I do not write a reference for a student unless I can be both positive and specific. It is therefore your responsibility to exemplify the type of student I can give a shining recommendation of, with characteristics such as: hard working, collegial, honest, and intellectually curious. If you are still early on in your college career, I encourage you to cultivate professional relationships with professors you might wish to ask a Letter of Recommendations from. Don't simply take a class, and then disappear until you would like a letter—your professor *will* forget your work and performance. If you are a senior, I encourage you to do your best to impress your professors with a stellar final year or semester.

IMPORTANT SAFETY INFORMATION

BCAL

If you have concerns about the safety or behavior of fellow students, TAs or Professors, call BCAL (the Behavior Concerns Advice Line): 512-232-5050. Your call can be anonymous. If something doesn't feel right – it probably isn't. Trust your instincts and share your concerns.

Evacuation Information

The following recommendations regarding emergency evacuation from the Office of Campus Safety and Security, 512-471-5767, <http://www.utexas.edu/safety/>

Occupants of buildings on The University of Texas at Austin campus are required to evacuate buildings when an alarm or alert is activated. Alarm activation or announcement requires exiting and assembling outside, unless told otherwise by an official representative.

- Familiarize yourself with all exit doors of each classroom and building you may occupy. Remember that the nearest exit door may not be the one you used when entering the building.
- Students requiring assistance in evacuation shall inform their instructor in writing during the first week of class.
- In the event of an evacuation, follow the instruction of faculty or class instructors. Do not re-enter a building unless given instructions by the following: Austin Fire Department, The University of Texas at Austin Police Department, or Fire Prevention Services office.

Link to information regarding emergency evacuation routes and emergency procedures can be found at: www.utexas.edu/emergency

Tentative 2018 Schedule

Week	Day	Date	Topic	Text Book Readings	Additional Readings	Assessment Due
0-1	Wed	29-Aug	Couse Into			
0-2	Fri	31-Aug	Lab0: Software Intro			
1-3	Mon	3-Sep	NO CLASS			
1-4	Wed	5-Sept	Pearson Correlation		Freund, Ch1 Bobko, Ch1	RQ 1
1-5	Fri	7-Sept	Lab1: Correlation			
2-6	Mon	10-Sept	Simple Linear Regression	Cannon, Ch1	Chatterjee, Ch2: 2.1 through 2.5	RQ 2
2-7	Wed	12-Sept	SLR Diagnostics			Lab 1
2-8	Fri	14-Sept	Lab2: Linear Regression I			
3-9	Mon	17-Sept	Inference around SLR	Cannon, Ch2	Chatterjee, Ch2: 2.6 through end	RQ 3
3-10	Wed	19-Sept	Confidence and Prediction			Lab 2
3-11	Fri	21-Sept	Lab3: Linear Regression II			
4-12	Mon	24-Sept	Additional Continuous Predictors; Centering revisited	Cannon, Ch3	Chatterjee, Ch3	RQ 4
4-13	Wed	26-Sept	Part Correlation and Multicollinearity			Lab 3
4-14	Fri	28-Sept	Lab4: Multiple Regression			
5-15	Mon	1-Oct	UNIT 1 Exam (wk0-wk4)			
5-16	Wed	3-Oct	Nesting regression models (and pitfalls)		Warner, R. Ch14	RQ 5; Lab 4
5-17	Fri	5-Oct	Lab5: Sequential Regression			
6-18	Mon	8-Oct	Coding of categorical predictors		Keith T., Ch6	RQ 6
6-19	Wed	10-Oct	Overall fit and pairwise tests			Lab 5
6-20	Fri	12-Oct	Lab6: Categorical Predictors			
7-21	Mon	15-Oct	The idea of moderation		Aiken & West, Ch7; Holmbeck, 2002	RQ 7
7-22	Wed	17-Oct	Simple slopes			Lab 6
7-23	Fri	19-Oct	Lab7: Categorical by Continuous Interactions			

8-24	Mon	22-Oct	Moderation continued		Aiken & West, Ch2 & Ch3; Hayes, 2009; Bauer & Curran, 2005	RQ 8
8-25	Wed	24-Oct	J-N technique			Lab 7
8-26	Fri	26-Oct	Lab8: Continuous by Continuous Interactions			
9-27	Mon	29-Oct	UNIT 2 Exam (wk5-wk9)			
9-28	Wed	31-Oct	Mediation and the Sobel test		Hayes, Ch4; Baron & Kenny, 1986	RQ 9;Lab 8
9-29	Fri	2-Nov	Lab9: Mediation Models			
10-30	Mon	5-Nov	Segmented Regression		Wagner, 2002; Trochim, 1990	RQ 10; Poster Idea
10-31	Wed	7-Nov	Interaction Segmentation			Lab 9
10-32	Fri	9-Nov	Lab10: Segmented Regression			
11-33	Mon	12-Nov	Dichotomous outcomes		Warner, Ch24	RQ 11
11-34	Wed	14-Nov	The logistic model			Lab 10
11-35	Fri	16-Nov	Lab11: Logistic Reg. (Odds)			
12-36	Mon	19-Nov	ROC and AUC		TBD	RQ 12
NA	Wed	21-Nov	Thanksgiving Break			
NA	Fri	23-Nov	Thanksgiving Break			
13-37	Mon	26-Nov	Multinomial logistic regression		TBD	
13-38	Wed	28-Nov	Ordinal Logistic			Lab 11
13-39	Fri	30-Nov	Lab 12: Multinomial Logistic Regression			
14-40	Mon	3-Dec	ANOVA and Regression		TBD	RQ 13
14-41	Wed	5-Dec	When Regression goes wrong.			Lab 12
14-42	Fri	7-Dec	Applied Extra Lab			
15-43	Mon	10-Dec	Final Unit Exam Day (wk10-wk14)			

FINAL (Poster Session): Monday, December 17, 9:00 am-12:00 pm