Advanced Methods for Statistical Inference
ITAO70200

Office: 337 Mendoza

Email: seth.berry@nd.edu

Office Hours

Monday & Wednesday: 11:00 to 1:30

Friday: 10:00 to 12:00

If you find my door ajar to any degree (it will typically be less than 10° and the office will be dark), then you are more than welcome to drop in and chat with me about absolutely anything (stats, programming, career, etc.).

Class Days and Time

Tuesday and Thursday – 3:00 to 4:50

Location – L068

Course Objectives

By the end of this class, I want for you to be able use the general linear model, generalized linear models, and interpret results from both. You will also be able to identify practical reasons for using mixed effects models and how Bayesian Statistics can be used. While mastery is not required, you will have a high-level understanding of these methods and you should be able to appropriately apply them to problems.

Attendance

Attendance in this course is not required, in that I will not be taking attendance; attendance is certainly recommended (and even encouraged). While the lecture presentations will be available, they are not verbatim recitations of what was covered and you may not rely on them to make it through everything in the course. Learning statistics takes effort and attending class is but one small part of that effort.

Although attendance is not required, we will have a weekly short comprehension check question. Each question is worth 10 points and is essentially a participation credit. These are

meant to be done in class – if you are not in class when they are given, you cannot submit them.

Readings

There is no official textbook for this course, but there are going to be a few assorted readings and resources for topics. I will also share resources for topics. We (read: you) are absolutely not committed to the readings; however, they will give you some very helpful background and I encourage you to read them; in other words, they are suggested readings, not required readings.

Homework

All homework assignments must be submitted in an html file knitted from Rmarkdown. The reasons for this requirement are mainly related to making sure that you have successfully gotten results from your code and that you are not just submitting code without running it first. Once you start using RMarkdown, it will be hard for you to go back to writing in a traditional word processor.

A significant portion of your grade will come from the homework map (i.e., your path towards statistical enlightenment). The skills roadmap has different levels (bronze, silver, and gold). Each level must be completed in order (i.e., you cannot skip silver and go straight to gold from bronze). Silver and gold levels essentially act as modifiers to not only your grade, but also your understanding.

Please feel free to work together on the bronze level of homework, but each assignment needs to be your own work. Putting your heads together to formulate an analytic attack plan is perfect (we all stands on the shoulders of giants), but copying and pasting text from each other is unacceptable. In other words, your code and words should not look like one of your classmate's – you will be graded accordingly.

Exam

There will be one exam during the course of this semester and it will be a takehome test. You **cannot** work together on this. You will have one week to complete the test.

Presentations

I promise you that these will be the easiest presentations you will ever give. Find data and prepare a 3 minute presentation (bonus points if you can explain it with one visualization!). These will be given during the designated final day and time. These will be graded on the content of your presentation (general creativity, appropriateness of analysis, etc.), not

necessarily on the scientific merit of your work! I really want you to be able to speak about your results in a defensible way.

Grade Breakdown

Comprehension Checks – 50 points (16%)

Homework – 90 points (29%)

Test – 120 points (39%)

Presentation – 50 points (16%)

Total – 310 points

Schedule

Table 1: Tentative Schedule

	Week	Date	Topic	Assignments
1	1	10/29 (T)	Introduction and R Markdown	
2		10/31 (TR)	Data Exploration and Visualization	
3	2	11/05 (T)	Categorical Data Analysis	Homework #1 Assigned
4		11/07 (TR)	The General Linear Model	
5	3	11/12 (T)	Sampling and Bootstraps	
6		11/14 (TR)	Regression Weirdness	Homework #2 Assigned
7	4	11/19 (T)	Generalized Linear Models	
8		11/21 (TR)	Generalized Linear Models	
9	5	11/26 (T)	Mixed Models	Midterm Assigned
10		11/28 (TR)	Thanksgiving	
11	6	12/03 (T)	Mixed Models	
12		12/05 (TR)	Bayesian Data Analysis	Homework #3 Assigned
13	7	12/10 (T)	Bayesian Data Analysis	
14		12/12 (TR)	The Inference/Prediction Divide	
15	8	12/17	Presentations (same time and place)	