

Research Design and Analysis I (16:830:521)
Fall 2019 (Last update: 9/4/2019)

Instructor Evan M. Kleiman, Ph.D.
Email: Evan.kleiman@rutgers.edu
Office hours: Tuesdays 2-3pm or by appointment
*Please email to confirm before coming to office hours. This will allow us to find an alternate meeting time if a large number of students are planning to attend.

Course time: Thursdays 9:30-12
Location: Tillett 102
Canvas site: <https://rutgers.instructure.com/courses/32524>

Texts

Required texts: Wickham, H., & Grolemund, G. (2016). *R for data science: Import, tidy, transform, visualize, and model data*. Sebastopol, CA: O'Reilly.

NOTE: You can buy this book, but the authors provide for free a better, easier-to-access version online at <https://r4ds.had.co.nz/data-visualisation.html>.

Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London: Sage.

NOTE: You don't need this until week 4. This should give you ample time to find a cheaper used copy if needed.

Other useful resources: R Studio Cheatsheets (<https://www.rstudio.com/resources/cheatsheets/>). Many of the skills we'll be learning in class have a corresponding cheatsheet that you can use.

Stack Overflow (<https://stackoverflow.com/>). You can find answers to many R questions.

Important class info

What is the structure of the class?

The first half (or so) of the class will be a didactic instruction where I will introduce and demonstrate the skills that we'll be learning. The second half will be a "hands on" chance for you to try out the skills taught in the first part of the class.

This course is a work in progress. What does that mean?

One of the most exciting things about Rutgers' graduate psychology programs is the diverse range of backgrounds of the students. One challenge with this diversity of experiences, however, is that students will come to this class with a very wide range of experience with statistics. This means that we need to adapt the class as we go so that everyone leaves with the same level of experience needed to succeed in a Ph.D. program.

What are the goals of this course?

My goal is for students to end this semester being: (1) Comfortable with the basics of the R statistical computing language. (2) Knowledgeable about the basics of inferential statistics. (3) Equipped to learn the other advanced statistics needed for a successful career as a researcher. This means that the goal of the class is not for you to learn every single statistical method that exists. This wouldn't be possible (because it would take years) and if it were possible, it wouldn't be useful since new methods are developed frequently.

But wait! There's more! The way that you'll be demonstrating to me that you've met these goals is by proposing a hypothesis using actual data (from your advisor or otherwise, see below), testing that hypothesis, and writing up the results. These steps will be spread across the semester and the requirements will grow as your knowledge grows. You should discuss this idea with your advisor after the first class. Ideally, the final paper you turn in to me will be something you could easily submit for publication with only a bit more work.

Do I need data for this course?

It would be ideal if you came to third week of class with a dataset that you are interested in and could possibly write a paper from. The dataset should have some demographic variables, a few continuous variables, and some categorical variables. If you have a dataset that doesn't meet all of these criteria, this is okay. We can generate other data for you to use for the in-class examples. If you can't get data, data will be provided to you.

Assignments

Out of class: Midterm and final (65% of total grade)

The midterm and final projects for this class will build upon each other. They will involve a writeup of an analysis you've done on your data or other data. We'll discuss more details well before each due date. There three main components:

Midterm/final proposal (Due 9/26, 5% of grade): This should be a brief (~2 paragraph max) writeup of the background of the hypothesis you want to test (1 paragraph) and the analysis you'll need to test it (1 paragraph). You should strive to propose analyses no more complicated than regression (so that the project fits within the scope of the class). But, I'm happy to help with more complex analyses, in some cases.

Midterm (Due 10/31, 25% of grade): This will involve a writeup of your "descriptive statistics" section of your midterm/final, including correlations, measures of central tendency, and related statistics. You'll have a clear sense (from me) about which specific statistics you'll need for your specific project as the project gets closer.

Final (Due 12/13; 35% of grade): This will involve a complete writeup of your project's method and results section. This will build on (and should include) what you turned in for your midterm.

In class: Participation and in-class assignments (35% of total grade)

As noted above, the class will involve a didactics portion and a hands-on portion. You'll demonstrate progress in these two areas in the following ways:

Participation (10% of grade). This reflects your attentiveness, during the lecture, the quality (but not quantity) of your contributions during the discussion, and being a good citizen of class (e.g., helping others, not missing classes).

In-class assignments (5 assignments, 5% each, 25% of grade total). Several of the labs will have hands-on activities that I will collect for points. You will usually be able to complete these before you leave class. More detail will be given early in the semester.

Course schedule

#	Date	Topic	What skills will I learn?	Reading	Due (before class)
1	9/5	Intro to the class, intro to R	How to install and launch R/Rstudio	None	
2	9/12	Data manipulation	How to work with datasets, subsetting, merging, basic descriptive statistics	RFDS Chapter 5 (skim)	
3	9/19				
4	9/26	Data visualization	The basics of visualizing data in <i>ggplot2</i> , the basics of variability and distributions	RFDS Chapter 3	
5	10/3	Correlation + effect sizes	Conducting correlations in R	DS Chapter 6	Midterm proposal
6	10/10	t-tests (and maybe chi-square tests)	t-test, chi-square tests (maybe)	DS Chapter 9 DS 18.1 - 18.6 (maybe)	
7	10/17	Regression I: The basics	OLS Regression in R	DS Chapter 7	
8	10/24	Regression II: Logistic regression	Logistic regression in R + related models (Poisson, negative binomial)	DS Chapter 8	
9	10/31	Regression III: Moderation (maybe)	Moderation in R (maybe)		Midterm
10	11/7	ANOVA	One-way ANOVA	DS Chapter 10	
11	11/14	Multi-level modeling I: The basics	A gentle intro to multi-level modeling in <i>lme4</i>	DS Chapter 19.1-19.6	
No class on 11/21: ABCT Conference					
12	11/26	Multi-level modeling II: Random slopes + growth models (maybe)	A slightly less gentle intro to multi-level modeling in <i>lme4</i>	DS Chapter 19.7	
13	12/5	Wrapping up / catching up			

FINAL DUE BY 11:59pm on 12/13

Other information

Attendance

This is a hands-on class and attending it is crucial to your success in class. However, the realities of being in graduate school mean that you will need to miss class from time-to-time for legitimate reasons (e.g., conferences). Please let me know in advance if you plan to miss a class. You will be responsible for learning the material you will miss during class. Document your absences using the absence reporting system: <https://sims.rutgers.edu/ssra/>

University Code of Student Conduct

It's important to realize that coming to the University brings you into a scholarly community, and as with all communities, there are principles and standards of behavior and action. Below, is the Preamble to the University Code of Student Conduct. (The full document can be found at <http://studentconduct.rutgers.edu/student-conduct-processes/university-code-of-student-conduct/>)