Emory University

RESEARCH METHODS & MODELS: STATISTICS

SOC 500; Fall 2019

Thursdays, 3:00 – 6:00, 206 Tarbutton Hall

Dr. Timothy J. Dowd
231 Tarbutton Hall
Office Hours: Mondays, 1:00 – 3:00
email: tdowd@emory.edu

Allison Roberts, M.A., R Support
Data Lab, 208 Tarbutton Hall
Office Hours: Wednesdays, 12:30 – 1:30
email: allison.roberts@emory.edu

COURSE DESCRIPTION

This course is an introduction to descriptive and inferential statistics for bivariate and multivariate analyses. The course will help you understand statistics reported in social science publications and in the news media, as well as help you conduct original research. The overall goal is to increase your statistical literacy – your ability to create, interpret, and critically evaluate statistical evidence. This is a set of skills that you will find highly useful in your current academic life and in your future career. It is also a valuable set of skills for virtually everyone in modern society, as statistical knowledge (and numerical literacy more broadly) is key for making sense of the growing amounts of information that we encounter in a digital world.

THE ORGANIZATION OF THIS COURSE

I approach this course with certain principles. First, I firmly believe that the use of statistics without a theoretical framework is highly problematic. Indeed, the best sociological scholarship weaves together theoretical concerns with empirical investigation. The language of the syllabus emphasizes the empirical aspects of scholarship, well, because this is a statistics course. But, please know that while we grapple with such things as data, sampling, regression – that is due to the nature of this course rather than my privileging of statistics over theory. (I also teach theory courses, by the way).

Second, I likewise do not privilege quantitative approaches over qualitative approaches. Both are extremely important for scholarship across the disciplines. The quantitative-heavy nature of this course is due to the fact, well, that this is a statistics course.

Third, I recognize that most practicing sociologists rely upon software packages when conducting statistical analyses (I sure do). As a result, this introductory course on statistics also is paired with a particular package – that of R. As several of the books emphasize, R has many advantages, but it also has a learning curve. We handle that curve in two ways. On the one hand, we ease into R by way of descriptive statistics and diagnostics before getting into data visualization and linear / logistic regression. Thus, this course picks up speed, so to speak, as you
become more adept at both statistics and the use of R. On the other hand, we also have an outstanding R support person that you can consult throughout the semester – Allison Roberts. You can see her contact information above.

Finally, I believe that the best way to master statistics is use them in the context of your own research (rather than simply doing the exercises found in a given text). I have set up this class, then, with the opportunity for us to engage in original research collectively. In particular, I have identified datasets (e.g., surveys) that allow us to explore our overlapping interests (see the weeks on creative workers), as well as to grapple with potentially new interests (see the week on the global circulation of music). My goal is that, rather than produce papers and exams for this course, you all will emerge with one or more papers suitable for journal submission in the very near future.

---

**COURSE REQUIREMENTS**

This course is a seminar, which entails, not only a fair amount of reading, but also class discussion and in-class exercises / analyses. Given that, you are expected to attend each class and to participate.

**A) Attendance**

If you must miss a class, please inform me ahead of time so that we can make arrangements. Note that unexcused absences will negatively affect your final grade.

**B) Class Participation and Discussion**

Active participation requires adequate preparation. You therefore must read the assigned material before class, access the appropriate datasets and codebooks for the week’s readings, and practice the skills necessary for statistical analysis and visualization.

Such careful preparation will improve the quality of class discussion. Of course, class discussion should be both informed and respectful; moreover, it should be a forum wherein all can raise questions, explore ideas, and express misgivings.

**C) Electronic Device Policy**

Electronic devices, such as laptops and tablets, can be very useful in terms of taking notes – and we will definitely rely upon them when conducting analyses in class. However, when students use them in class to do other things (e.g., surf the Internet, order from online stores, check social media posts), that proves detrimental both to student participation and discussion. Note that using devices in a non-curricular way will, most definitely, affect your final grade in a negative fashion.
D) Assignments

You will not be doing assignments for this class in the traditional sense. Of course, you are still expected to develop the necessary knowledge and skills. Indeed, in each of our classes, you will be demonstrating your mastery of the week’s materials. So come prepared!

The actual assignments, instead, will involve work on collective (or individually-authored) papers. Prior to the weeks of October 3rd, November 7th, and November 21st – I will circulate information on the analyses that we will be doing when approaching both creative workers and the global circulation of music.

E) Final Grade

My starting assumption is that you will earn an “A” in this seminar. If you do the work and contribute to the class, then all is fine. If you fall below expectations, I will let you know, so that you can adjust accordingly. Of course, if you are having difficulties, please contact me as well as Allison Roberts.

COURSE RESOURCES

We will be drawing upon five books for the bulk of our readings, books that I have provided for the class. They are as follows:


Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2017. *An Introduction to Statistical Learning with Applications in R*. Eighth printing. New York: Springer.


We will not read the Gareth et al. book in its entirety. That is because SOC 500 gets you to linear and logistic regression but, then, the next course, Applied Regression course (SOC 560), offers a semester’s long focus on regression. Nonetheless, the Gareth et al. book will remain a great resource.
We will read the occasional article or chapter in addition to our five books. Those will be posted on our Canvas site for SOC 500 (see http://canvas.emory.edu).

We will also direct our analytical attention to datasets located in our shared folder on Emory Box (see https://emory.account.box.com/login).

If you have any special needs, please contact me at the beginning of the semester, and we will discuss the necessary arrangements (for additional information, visit the Emory Office of Accessibility Services website; see http://equityandinclusion.emory.edu/access/index.html).

---

**COURSE SCHEDULE**  
*(Subject to Revision)*

**NOTE:** For each week of readings, please ensure that you have accessed the relevant *R* materials, as well as the datasets / codebooks mentioned in those readings.

**August 29: Introductions**

**September 5: Situating Data & Statistics (and Getting Started on R)**


*Data Visualization: A Practical Introduction.* Preface and Chapter One.

*Bit by Bit: Social Research in the Digital Age.* Preface, Chapters One and Two.

*Basic Statistics for the Behavioral and Social Scientists Using R.* Appendix C and Chapter One.

**September 12: From Gathering Data to Descriptive Statistics**

*Bit by Bit: Social Research in the Digital Age.* Chapter Three.

*Basic Statistics for the Behavioral and Social Scientists Using R.* Chapters Two through Four.

**September 19: More Ways of Gathering Data, as well as Inferential Statistics**

*Bit by Bit: Social Research in the Digital Age.* Chapter Four.

*Basic Statistics for the Behavioral and Social Scientists Using R.* Chapters Five through Seven.
September 26: Depicting Patterns Amidst All That Data

Data Visualization: A Practical Introduction. Chapters Two through Five.

October 3: Describing and Depicting the Situation of Creative Workers


October 10: Considering Online Data Opportunities and Ethics

Bit by Bit: Social Research in the Digital Age. Chapters Five and Six.

October 17: An Initial Foray into Linear Regression and Its Visualization

Data Visualization: A Practical Introduction. Chapter Six.

Basic Statistics for the Behavioral and Social Scientists Using R. Chapters Eight through Ten.

October 24: More on Linear Regression


An Introduction to Statistical Learning with Applications in R. Chapters One through Three.

October 31: Moving on to Logistic Regression


An Introduction to Statistical Learning with Applications in R. Chapter Four.

November 7: Using Regression to Consider the Situation of Artists and Musicians


**November 14: Considering the Spatial Patterns Amidst All That Data**

*Data Visualization: A Practical Introduction*. Chapters Seven and Eight.

**November 21: Mapping the Global Circulation of Music**


**November 28: THANKSGIVING BREAK (No Class)**

**December 5: Looking Back, and Looking Ahead**


*Bit by Bit: Social Research in the Digital Age*. Chapter Seven.