Advanced Statistical Methods
POL 518/618
Tuesdays and Thursdays 5:05-6:20pm, Dooly Memorial 101
Spring 2020

Professor Calla Hummel
Dr./she/her
Office hours: Tuesdays and Wednesdays 3:30-4:30pm
Office: 1300 Campo Sano, 160F
Email: chummel@miami.edu

Course Description:
This class teaches students advanced mathematical and computational methods for analyzing quantitative social science data. The course covers data cleaning, multiple regression with continuous and binary variables, and basic statistical coding, as well as writing and publishing in political science. Class time will include statistical lectures as well as hands-on data labs and coding exercises. Students will develop a group final project where they answer a social science question with a statistical analysis and interpret the analysis in a social science paper. Special topics lectures will depend on the needs of student research projects and can include time series, machine learning, or panel data. The course uses the open source R software.

Course Requirements: Laptop

Course Prerequisites: POL 517, POL 610 or other introductory statistics course

Required books (from the Sage Greenback Quantitative Applications in the Social Sciences series):
- Lewis-Beck, Colin, and Michael Lewis-Beck. *Applied Regression: An Introduction*. Vol. 22. Sage Publications, 2015.
- Liao, Tim Futing. *Interpreting Probability Models: Logit, Probit, and Other Generalized Linear Models*. No. 101. Sage Publications, 1994.

Additionally, I will regularly assign published political science papers that develop or use the methods we discuss.

Recommended review in pictures (I do not require these because they cover introductory statistics concepts and because the gender tropes are a bit outdated; however, some students have found the visuals very useful):
- Gonick, Larry, and Woollcott Smith. *The cartoon guide to statistics*. Collins, 1993.
- Takahashi, Shin. *The manga guide to statistics*. No Starch Press, 2008.

Recommended reference texts (the first three are standard econometric texts for those who want to continue learning on their own. The last two are available for free and are

standard statistical learning texts. Note: not that much changes between editions in the econometrics books):

- Wooldridge, Jeffrey M. *Introductory econometrics: A modern approach*. Nelson Education, 2015.
- Gujarati, Damodar N., and Dawn C. Porter. "Essentials of econometrics." (1999).
- Kmenta, Jan. "Elements of econometrics." *Ann Arbor: MI: University of Michigan Press* (1986).
- Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. *The Elements of Statistical Learning*. New York: Springer series in statistics, 2001. 12th edition free at: https://web.stanford.edu/~hastie/ElemStatLearn/
- StatSoft. *Electronic Statistics Textbook.* 2017. Free at: http://www.statsoft.com/Textbook

Evaluation:
Evaluation occurs throughout the semester through assignments that build up to the final project. Additionally, students will be graded on class participation and their final project. Every assignment is worth points and grades are out of 100 possible points:

• Class Participation (10%): Every student is expected to show up to every class prepared to ask questions and solve problems.

• R Tutorial (10%): Show proof of completion of an online R tutorial.

• Prospectus (10%): Outline and develop an idea for a final project.

• Midterm (20%): There will be a take home midterm that will test your ability to locate data, load it into R, and analyze it.

• Final Project Draft (15%): Students submit a draft of their group final paper two weeks before the final paper is due.

• Final Project (35%): Students submit a final statistical paper as a group. The paper should present a research question, an argument, a literature review, a research design, a statistical analysis of that data, and a discussion of findings, as well as replication data and code. The paper should be at least 6,000 words and at least 10 works cited (ideally many more). The document should be double-spaced, 12 point font, with one inch margins, black ink, and in Times New Roman, Arial, or Cambria font.

• Alternative final project (50%): Any student who does not want to do the class project can work as an RA on one of my statistical projects. I will assign RAs weekly and bi-weekly data collection, cleaning, and appendix writing tasks.

There are no make up assignments. Late assignments will be docked 20% for every

day late.

Grades:
At the end of the semester, the instructor will convert points to a letter grade on the following scale:
93-100 = A
90-92 = A-
87-89 = B+
83-86 = B
80-82 = B-
77-79 = C+
73-76 = C
70-72 = C-
67-69 = D+
63-66 = D
60-62 = D-
59 or less = F

Grade disputes:
If a student believes a grade was issued in error, they must submit an explanation of the error in writing within two days of receiving a graded assignment. The student must then meet with the instructor to discuss the problem. If the instructor agrees to re-grade the assignment, the new, final grade may be higher or lower than the original.

Email policy:
I will only answer questions in person, not over email. If you email me in a non-emergency, I will not respond. You will get more help and attention if you talk to me in person. I will only respond to emails setting up a time to meet me or alerting me to an emergency that prevents you from talking to me in person. I have office hours twice a week and if you cannot make those, I am happy to set up an in-person meeting at another time.

Expectations:
- **Absolutely no phones during class and laptops are for note-taking and analysis only**
- You are expected to take thorough notes.
- Show up to class on time and do not pack up or leave until class is over.
- Do the reading before the class.
- Come to every class (unless you are sick or have an excused absence). Students are graded on class participation.
- Stay home if you are sick. You do not need to email me unless you will have an extended absence.

Names, pronouns, and self-identification:
We recognize the importance of a diverse student body, and we are committed to fostering equitable classroom environments. I invite you, if you wish, to tell us how you

want to be referred to both in terms of your name and your pronouns (he/him, she/her, they/them, etc.).

Additionally, how you identify in terms of your gender, race, class, sexuality, religion, and dis/ability, among all aspects of your identity, is your choice whether to disclose (e.g. should it come up in classroom conversation about our experiences and perspectives) and should be self-identified, not presumed or imposed.

I will address and refer to all students accordingly and will support you in doing so as well. I expect students to use and respect their classmates' stated names, pronouns, and identities.

Religious holidays and other excused absences:
You must notify me of your pending absence by the third class session. If you must miss a class, an examination, a work assignment, or a project in order to observe a religious holy day or other permitted absence, you will be given an opportunity to complete the missed work within 7 days after the absence.

Students with disabilities:
If you have a disability concern and/or accommodation, talk to me the first week of class in office hours or over email. If you haven't already, familiarize yourself with the services available through the Camner Center: http://camnercenter.miami.edu

Plagiarism:
I will fail you if you plagiarize. Plagiarism refers to using of other peoples' written work (from books, articles, newspapers, web sources, etc.) or verbal ideas without attribution. You must cite all material that is not your own. For example, while I encourage you to discuss class readings with your peers, you must write reading summaries yourself. If you have doubts of what constitutes plagiarism, consult the student code, and/or talk to me. "I did not know" is not an acceptable defense after a work has been handed in.

For any additional concerns, please do not hesitate to ask me. If you anticipate or begin to experience difficulties in the class, I recommend that you come talk to me as early as possible. My office hours are set up for you, and I encourage you to use them.

Schedule

Take-home midterm tentatively on 2/25
Prospectus due 3/17
Draft group paper due 4/7
Final group paper due 4/30

Topics we will cover sequentially:
- Introduction to R and R Studio
  - Download R: https://www.r-project.org

- o Download R Studio (free version): https://www.rstudio.com/products/rstudio/download/
- o Read: R-Bloggers's Introduction to R https://www.r-bloggers.com/how-to-learn-r-2/#h.nl3rvdr3ycsc
- o **Do (for a grade):** https://www.r-bloggers.com/learning-r-the-ultimate-introduction-incl-machine-learning/
- o Recommended: There are a ton of more in-depth courses on R programming and R programming textbooks. Many of the paid courses have free 10-day trials and you can complete an R course in that time.

- Why Do We Care about Statistical Inference and Data Science?
  - o Stauffer, Katelyn E., and Diana Z. O'Brien. "Quantitative Methods and Feminist Political Science." *Oxford Research Encyclopedia of Politics*. 2018.
  - o D'Ignazio Catherine, and Lauren Klein. "Chapter One: Bring Back the Bodies." In *Data Feminism*. Online: https://bookbook.pubpub.org/pub/zrlj0jqb?readingCollection=09555901
  - o Recommended: Cathy O'Neil. *Weapons of Math Destruction*.
  - o Recommended: Caroline Criado-Perez: *Invisible Women*.

- What Can We Do with a Dataset?
  - o Alex Hanna on Bad Hessian: RuPaul's Drag Race Data Tutorial. http://badhessian.org/2013/03/lipsyncing-for-your-life-a-survival-analysis-of-rupauls-drag-race/
  - o Read one of these:
    - ▪ Hummel et al working paper on LGBT rights
    - ▪ Hummel, Calla, John Gerring, and Thomas Burt. 2019. "Do Political Finance Regulations Reduce Corruption?" *British Journal of Political Science* (Panel analysis example; replication data and code included.)

- Probability, Sampling, and Hypothesis Testing
  - o Review a cartoon statistics guide if you bought one
  - o Hicks, Mar. "Hacking the Cis-tem." *IEEE Annals of the History of Computing* 41.1 (2019): 20-33.
  - o Recommended: Hicks, Mar. *Programming Inequality*. MIT Press, 2017.

- Variance, Covariance, and Correlation
  - o Chapter 1 and 2 of Lewis-Beck and Lewis-Beck

- Descriptive Statistics and Handy Graphics in R
  - o Holland, Alisha C. "Forbearance." *American Political Science Review* 110.2 (2016): 232-246.

- Ordinary Least Squares Estimation and Its Assumptions

- Chapter 3 and 4 of Lewis-Beck and Lewis-Beck

- Dichotomous Dependent Variables: Probit and Logit
  - Chapter 1, 2, and 3 of Liao
  - Philpot, Tasha S., and Hanes Walton Jr. "One of our own: Black female candidates and the voters who support them." *American Journal of Political Science* 51.1 (2007): 49-62.

- Data Management and Data Cleaning in R

- Multivariate Research Design
  - Acharya, Avidit, Matthew Blackwell, and Maya Sen. "The political legacy of American slavery." *The Journal of Politics* 78.3 (2016): 621-641.

Special topics that we will cover if we have time and if they are useful:
- Time-Series Cross-Sectional or Panel Analyses
  - Beck, Nathaniel, and Jonathan N. Katz. "What to do (and not to do) with time-series cross-section data." *American Political Science Review* 89.3 (1995): 634-647.
  - Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. "Taking time seriously: Time-series-cross-section analysis with a binary dependent variable." *American Journal of Political Science* (1998): 1260-1288.
- Statistical Learning/Machine Learning
  - Muchlinski, David, et al. "Comparing random forest with logistic regression for predicting class-imbalanced civil war onset data." *Political Analysis* 24.1 (2015): 87-103.
  - Roberts, Margaret E., et al. "Topic models for open ended survey responses with applications to experiments." *American Journal of Political Science* (2014).
- Matching Case Studies and Statistical Analysis
  - Fairfield, Tasha, and Andrew E. Charman. "Explicit Bayesian analysis for process tracing: guidelines, opportunities, and caveats." *Political Analysis* 25.3 (2017): 363-380.
- Writing and Publishing in Political Science
  - Gerring, John. ND. "General Advice on Social Science Writing."
  - King, Gary. "Publication, publication." *PS: Political Science & Politics* 39.1 (2006): 119-125. (with online notes: https://gking.harvard.edu/papers)