POL 304H1-S: Using Data to Understand Politics and Society
WINTER 2021

PREREQUISITES: POL232H1/POL242Y1/POL322H1/equivalent

Lecture: Monday, 5–7 pm, online–synchronous
Instructor: Professor Olga Chyzh, olga.chyzh@utoronto.ca
Office Hours: by appointment, `https://chyzh.youcanbook.me/`

Overview and Objectives

Students will learn about role of data for understanding politics and society, and a wide range of approaches to analyzing such data. How do information and data shape politics and policies? What insights can we gain about contemporary and past societies from the data they produce? To answer these questions we will examine a series of methodological approaches to different types of political data, including text analysis, network analysis, spatial statistics, and time-series analysis. This course will draw from topics in the study of international relations, economics, political behavior, and statistics to offer a diverse set of tools for processing and analyzing different types of data. Applications will include war and conflict, terrorism, international trade, social media, elections, and representation.

Learning Outcomes

This course is designed as a series of weekly modules that loosely build upon each other. Each module covers one or more state-of-the-art approaches to statistical data analysis. For each topic covered, the objectives are that students will:

- Learn the general mechanics of the specific method;
- Formulate theories and derive hypotheses testable using this method;
- Apply the method to extract/analyze real-world political and social data.

Course Materials

Materials for the course, including course videos, are posted on Quercus. Please allow time for processing videos.

Software

R (latest version) `https://www.r-project.org/`
RStudio (latest version) `https://rstudio.com/products/rstudio/download/`

Reading

Please complete all assigned readings prior to class.

Coding Sessions

For each method covered we will run through applications in R during class. Students are strongly encouraged to follow along during class and review/run through these examples after class. Students will be provided with data, but may also use their own datasets.

Grading Scale

Students will demonstrate their mastery of the course materials by completing three mini-projects (each making up 15% of the final grade) and a final project (40%). The remaining 15% will come from participation. If you are unable to attend/participate in the class discussion synchronously, please contact the instructor for an alternative assignment **during the first week of class** (a failure to do so will result in a score of 0 on the participation component). No late assignments are accepted. Students who are experiencing extenuating circumstances that may prevent them from completing an assignment should contact the instructor as soon as possible. The final grade will be calculated using the following grading scheme.

A+	≥ 90
A	≥ 85
A-	≥ 80
B+	≥ 77
B	≥ 73
B-	≥ 70
C+	≥ 67
C	≥ 63
C-	≥ 60
D+	≥ 57
D	≥ 53
D-	≥ 50
F	< 50

Course Policies

Student Responsibilities in the Learning Process: Students are expected to complete any assigned readings prior to completing that topic's assessment. Students are also expected to complete all assessments on time. This means accessing the materials with sufficient time to complete assessments prior to deadlines. In the event that a student has questions concerning the material, they should formulate specific questions to ask the professor via office hours or email with sufficient time for a response prior to assessment deadlines (i.e. emailed

questions should be sent at least 24 hours prior to a deadline, excluding weekends).

Classroom Conduct: Students are expected to participate in class in a thoughtful and respectful manner while in the pursuit of knowledge accumulation. Generally, this means engaging with one another's ideas and treating others as you would like to be treating as well as *not* treating others how you would *not* like to be treated. Please see university policies on freedom of speech (`https://governingcouncil.utoronto.ca/secretariat/policies/freedom-speech-statement-may-28-1992`) and discrimination and harassment (`https://governingcouncil.utoronto.ca/secretariat/policies/harassment-statement-prohibited-discrimination-and-discriminatory-harassment`).

Accommodations: Please discuss any special needs with the instructor start of the semester, for example, to request reasonable accommodations if an academic requirement conflicts with your religious practices and/or observances. Those seeking accommodations based on disabilities should complete the appropriate documentation with Student Life Programs and Services (`https://studentlife.utoronto.ca/department/accessibility-services/`).

Academic Misconduct: All acts of dishonesty in any work constitute academic misconduct. The Student Disciplinary Regulations (`https://governingcouncil.utoronto.ca/secretariat/policies/code-behaviour-academic-matters-july-1-2019`) will be followed in the event of academic misconduct.
A special note on plagiarism: plagiarism is the act of representing directly or indirectly another person's work as your own. It can involve presenting someone's speech, wholly or partially, as your; quoting without acknowledging the true source of the quoted material; copying and handing in another person's work with your name on it; and similar infractions. Even indirect quotations, paraphrasing, etc., can be considered plagiarism unless sources are properly cited.

Copyright: Course materials, including recorded lectures and slides, are the instructor's intellectual property covered by the Copyright Act, RSC 1985, c C-42. Course materials posted on Quercus or the class website may not be posted to other websites or media without the express permission of the instructor. Unauthorized reproduction, copying, or use of online recordings will constitute copyright infringement.

Course Schedule

Week 1: Introduction to R, Rstudio, and RMarkdown

Week 2: Text Analysis

- Jodie Archer and Matthew L. Jockers. *The Bestseller Code: Anatomy of the Blockbuster Novel*. St. Martin's Publishing Group, 2016. Chapter 4.

Week 3: Web-scraping

- Mini-Project 1 is due on Friday, January 29, at 5 pm.

Week 4: Big Data

- Danah Boyd and Kate Crawford. Six provocations for big data. a decade in internet time. Symposium on the Dynamics of the Internet and Society., 2011.
- David Lazer et al. The parable of Google flu: Traps in big data analysis. 2014.
- Aylin Caliskan, Joanna J. Bryson, and Arvind Narayanan. Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334):183–186, 2017.
- Jevin D. West. How to improve the use of metrics: Learn from game theory. *Nature*, 465(17):870–872, 2010.

Week 5: Networks I: Theory

- Mark S. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(6):l360–1980, 1973, **Trigger Warning:** This article uses racial and gendered language common at the time of its writing.
- John F. Padgett and Christopher K. Ansell. Robust action and the rise of the Medici, 1400-1434. *American Journal of Sociology*, 98(6):1259–1319, 1993.
- Apply for a Twitter developer account: `https://developer.twitter.com/en`. Click "Apply" in the top right corner and fill out the information to the best of your ability.

Week 6: Networks II: Social Media Applications

- Mini-Project 2 is due on Friday, February 26, at 5 pm.

Week 7: Spatial Analysis I: Maps

Week 8: Spatial Analysis II: Modeling Spatial Dependence

- Nathaniel Beck, Kristian Skrede Gleditsch, and Kyle Beardsley. Space is more than geography: Using spatial econometrics in the study of political economy. *International Studies Quarterly*, 50(1):27–44, 2006.

Week 9: Analysis of Temporal Data

- Donald T Campbell and H Laurence Ross. The Connecticut crackdown on speeding: Time-series data in quasi-experimental analysis. *Law and Society Review*, pages 33–53, 1968.
- Michael S. Lewis-Beck. Interrupted time series. In William D. Berry and Michael S. Lewis-Beck, editors, *New tools for social scientists: Advances and applications in research methods*. Beverly Hills, CA: Sage, 1986. Chapter 9.
- Olga V Chyzh and Robert Urbatsch. Bean counters: The effect of soy tariffs on change in republican vote share between the 2016 and 2018 elections. *The Journal of Politics*, 83(1):000–000, 2021
- Mini-Project 3 is due on Friday, March 19, at 5 pm.

Week 10: Survival Analysis

- Jason T. Rich, J. Gail Neely, and D. Phil Randal C. Paniello, and Courtney C. J. Voelker, Brian Nussenbaum, and Eric W. Wang. A practical guide to understanding kaplan-meier curves. *Otolaryngology–Head and Neck Surgery*, 143(3):331–336, 2010.

Week 11: Experiments

- Timothy D Wilson, Elliot Aronson, and Kevin Carlsmith. The art of laboratory experimentation. 2010.
- Joshua Aronson, Michael J Lustina, Catherine Good, Kelli Keough, Claude M Steele, and Joseph Brown. When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of experimental social psychology*, 35(1):29–46, 1999.

Week 12: Instrumental Variables

- Matthew Potoski and R Urbatsch. Entertainment and the opportunity cost of civic participation: Monday night football game quality suppresses turnout in US elections. *The Journal of Politics*, 79(2):424–438, 2017.
- Emily Hencken Ritter and Courtenay R Conrad. Preventing and responding to dissent: The observational challenges of explaining strategic repression. *American Political Science Review*, 110(1):85–99, 2016.

- Final Project is due on Friday, April 9, at 5 pm.