Informatics Lecture Room

Election Data Science

Prof. Michael McDonald

<u>Contact Info</u> <u>Office Hours</u>

Office: Anderson 222 W 9:30am-11:30am or by appointment

E-mail: michael.mcdonald@ufl.edu

Phone: 352-273-2371

Course Overview

Campaign consultants since the late 1960s analyzed and visualized large-scale databases in the areas of redistricting, campaign finance, and voting. The latter was popularized by the Obama campaign's large-scale microtargeting efforts in 2008. Recent years have seen yet another transformation of election consulting and journalism with the advent of analysis and visualization of election data. Popularized by organizations like 538, data analytics have become a routine part of how the public consumes election information.

Course Objectives

The goals of this class is to learn the properties of election data, how to analyze it, and how to visualize it. While we focus on election data, job applicants who have the analytics skills we will cover in this class are in high demand. We will primarily use a programming language called R. If you don't already have it, it is freely available at https://www.r-project.org/. You may also wish to download a popular program that provides a user interface called R Studio, which is also freely available at https://www.rstudio.com/. Along the way to learning how to use R, we will apply programing tasks to real-world election examples.

Alumni of this class have provided their class project as work product examples, which have helped them in securing a job upon graduation. I expect that the analysis and visualization skills you learn here will assist you in your future careers in politics or data science.

Course Materials

- Hadley Wickham and Garrett Grolemund, 2nd edition. 2023. <u>R for Data Science</u>. Sebastopol, CA: O'Reilly Media Inc.
- Kieran Healy. 2018. <u>Data Visualization: A Practical Introduction</u>. Princeton University Press.
- Additional readings are available online on Canvas.

You do not need to purchase a textbook! I find having a hard copy sometimes useful, but having one is not necessary. Free online versions of the textbooks are available through the links provided.

The readings that will be discussed each class are listed below as a <u>Reading Assignment</u> for that day of class.

Additional Academic Resources

- <u>Career Connections Center</u>: Reitz Union Suite 1300, 352-392-1601. Career assistance and counseling services.
- <u>Library Support</u>: Various ways to receive assistance with respect to using the libraries or finding resources.
- <u>Teaching Center</u>: Broward Hall, 352-392-2010 or to make an appointment 352-392-6420. General study skills and tutoring.
- <u>Writing Studio</u>: 2215 Turlington Hall, 352-846-1138. Help brainstorming, formatting, and writing papers.
- Student Complaints On-Campus: <u>Visit the Student Honor Code and Student Conduct Code webpage for more information.</u>
- On-Line Students Complaints: <u>View the Distance Learning Student Complaint Process</u>

Enrollment Management Complaints (Registrar, Financial Aid, Admissions): <u>View the Student Complaint Procedure webpage for more information.</u>

Attendance Policy

If you do not participate in at least one of the first two class meetings of a course or laboratory in which you are registered, and you have not contacted the professor to indicate your intent, you can be dropped from the course. You must not assume that you will be dropped, however. The department will not notify you if you have been dropped from a course or laboratory. You can request reinstatement on a space-available basis if you present documented evidence of your absences.

The university recognizes the right of the individual professor to make attendance mandatory. After due warning, professors can prohibit further attendance and subsequently assign a failing grade for excessive absences.

For more information on the university's attendance policies, see here.

In my experience, students who fail to attend class tend to have a personal problem affecting their performance. If you're struggling, please contact the instructor so we can devise a plan together. The university provides <u>resources to students in distress</u> that I encourage you to seek out.

The policy for this class is that every student should make the best attempt to attend all classes. A student that misses four or more classes without a university-approved excuse (e.g., from the Dean's office, sports department, or other university official) will

have their overall grade reduced a letter step (e.g., B+ to B). Students do not need, and should not approach the professor, to provide non-university-approved excuses for missing three or fewer classes. A student with ten or more unexcused class absences will receive a failing grade for the entire course and will be barred from attending further classes.

If you are late to class the instructor may not credit your presence as attending. This penalty is usually invoked for chronic tardiness and is usually preceded by a warning.

If there is an issue that affects your ability to attend class, please let the professor know as soon as possible. Accommodation can be made if we devise a plan together as soon as a problem occurs.

Grades

Grades consist of a class project and workshops to track your progress. The weighting of the aspects of the grade are:

<u>Assignment</u>	<u>Due</u>	% of Overall Grade	
Class Project			
Proposal	Sept. 26	10%	
Final Submission	Dec. 9	80%	
Workshops	Various	10%	

Class Project Proposal (10% of overall grade): A two-page double-spaced summary of a student's proposed class project. The proposal accounts for 10% of the class grade. The primary purpose of the proposal is to ensure students devise a manageable class project. Students may be required to amend their proposal based on the feasibility of the proposed project.

Class Project (60% of overall grade): The final project is a website that includes an introduction, data analysis visualizations from techniques learned in class, and sample code. Students are encouraged to discuss challenges they encountered and their solutions.

Class Assignments

Proposal

There are two paths for class projects, a research project and a campaign simulation.

Research projects are most appropriate for students who plan to conduct research in their future careers. You will write a one or two-page memo outlining your intended project. You must identify:

- 1. Why this is an interesting topic/guestion
- 2. The data you will analyze

3. What you expect to discover (i.e., a hypothesis)

Research proposals are graded on:

- (40% of proposal grade) Why their topic is interesting and what they expect to discover (i.e., a question framed as a hypothesis)
- (30% of proposal grade) The data that will be analyzed and possible analysis approaches.
- (30% of proposal grade) professionalism of the proposal, e.g., spelling, grammar, and formatting.

The campaign simulation path (discussed below) is most appropriate for students who wish to work on campaigns, particularly students of UF's Masters in Campaigning. Students should discuss their desired selection with the instructor for suitability before submitting their desired district or locality.

Campaign simulation proposals are graded on:

- (10% of proposal grade) Suitability of the district or locality
- (60% of proposal grade) Links to data needed for this project. (Voter file data may be provided by the instructor if it is not freely available from a public portal. You are still required to identify where you could obtain these data if you needed it for your campaign.)
- (30% of proposal grade) professionalism of the proposal, e.g., spelling, grammar, and formatting.

Class Project

There are two paths for graduate students' class assignments.

The first path is for students who wish to conduct independent research.

You will create a webpage describing your project, data visualization(s), code snippets and a description how you created your visualization(s), and a description of what the visualizations tell us. The webpage can be generated using R Markdown and publish to RPubs (https://rpubs.com/about/getting-started). You may publish to another platform if you wish to do so.

All students' class projects must be an original analysis of election administration data. Election administration data are data such as voter registration files, district or precinct boundary files, and precinct election results. These data <u>do not</u> include survey data, and projects analyzing survey data are inappropriate for the class project.

Student may conduct original data collection, for example, by contacting election offices. If you intend to such independent data collection please discuss your topic with the instructor as soon as possible.

Where appropriate to the research question and data, conduct advanced statistical analyses, such as difference-in-difference or ecological inference techniques in addition to more common statistical models.

The following elements are required for an independent research project:

- (25% of project grade) Overall narrative, which should incorporate text from the class project proposal. The narrative should explain how chosen visualizations address the research question. The narrative should include a discussion of findings and possible future research at the end.
- (50% of project grade) Students are graded on the appropriateness of their visualizations to their research project and class content, a description of the code required to generate the visualization, and any challenges that needed to be overcome. The types of required visualizations depend upon the student's research question and data. Typically, students are required to create a mapping visualization and line or bar charts.
- (25% of project grade) professionalism of the proposal, e.g., spelling, grammar, and formatting of text; clarity of sample code; labeling and aesthetics of visualizations.

Workshops

Throughout the semester, we will have workshops typically on Thursday classes. You will be given an assignment on Tuesday to complete by Thursday. These Thursday classes will be days where we collectively work through the problem in class. The assignment will be due that Thursday evening. Grades will be pass/fail for successful completion. Meaning, you can receive a grade of 100 or 0 for each completed assignment. A partially completed assignment counts as a fail or 0. There are six scheduled workshop assignments. You will be allowed to drop one from your grade calculation.

Grading Scale

Your number grade on assignments is converted to a letter grade using the following scale:

Percentage Earned	Letter Grade
93-100	Α
90-92	A-
87-89	B+
83-86	В
80-82	B-
77-79	C+
73-76	С
70-72	C-
67-69	D+
63-66	D
60-62	D-

Below 60	E

Your grade can be adjusted further for failure to adhere to the attendance and technology use policies.

Please be aware that a C- is not an acceptable grade for graduate students. The GPA for graduate students must be 3.0 based on 5000 level courses and above to graduate. A grade of C counts toward a graduate degree only if based on credits in courses numbered 5000 or higher that have been earned with a B+ or higher.

Your class grade translates into university grade points. For more information please refer to <u>university policies</u>.

Policy Related to Make Up Exams or Other Work

Students are required to submit work when due. The due dates for workshops are the midnight of the assigned date. In most cases this provides students an opportunity to receive feedback in class before making a final submission. You are strongly encouraged to work on your workshop assignment before it is covered in class. Keep in mind that you may drop one workshop from your grade if you cannot complete it for any reason.

If students cannot complete assignments on the due date, they must notify the instructor in advance that they will be unable to do so. A proper remedy will be discussed given the circumstances of the tardiness. University-approved excuses, such as circumstances verified by the Dean of Students or athletic travel, typically receive no late grade penalty if the assignment is promptly completed by a new due date. Other failures may result in grade reductions or a failing grade.

Technology Use Policy

Laptops are welcome in the classroom since there is a programming element to the course. However electronic devices are to be used only for class-related activities. **Don't use cell phones and otherwise browse devices or laptops for non-class content.** A first offense of using electronic devices for other than approved uses earns a warning. A second offense results in a full letter grade reduction of the overall class grade, and a third offense results in an automatic failing grade for the course. **Recordings**

Students are allowed to record video or audio of class lectures. However, the purposes for which these recordings may be used are strictly controlled. The only allowable purposes are (1) for personal educational use, (2) in connection with a complaint to the university, or (3) as evidence in, or in preparation for, a criminal or civil proceeding. All other purposes are prohibited.

Specifically, students may not publish recorded lectures without the written consent of the instructor.

A "class lecture" is an educational presentation intended to inform or teach enrolled students about a particular subject, including any instructor-led discussions that form part of the presentation, and delivered by any instructor hired or appointed by the University, or by a quest instructor, as part of a University of Florida course. A class lecture does not include lab sessions, student presentations, clinical presentations such as patient history, academic exercises involving solely student participation, assessments (quizzes, tests, exams), field trips, private conversations between students in the class or between a student and the faculty or lecturer during a class session. Publication without permission of the instructor is prohibited. To "publish" means to share, transmit, circulate, distribute, or provide access to a recording, regardless of format or medium, to another person (or persons), including but not limited to another student within the same class section. Additionally, a recording, or transcript of a recording, is considered published if it is posted on or uploaded to, in whole or in part, any media platform, including but not limited to social media, book, magazine, newspaper, leaflet, or third party note/tutoring services. A student who publishes a recording without written consent may be subject to a civil cause of action instituted by a person injured by the publication and/or discipline under UF Regulation 4.040 Student Honor Code and Student Conduct Code.

Honor Code

UF students are bound by The Honor Pledge which states, "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment." The Honor Code (http://www.dso.ufl.edu/sccr/process/student-conduct-honor-code/) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor.

Basically, don't cheat. You cheat yourself of your education and more severe punishment may follow. Assisting someone else complete a project is cheating, too. If you have any questions if your approach to completing an assignment may violate the honor code, please contact the instructor for guidance. This is true for any class.

Disability Statement

Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, www.dso.ufl.edu/drc/) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester and present their accommodation letter to the instructor so we can devise a plan for your success.

Counseling and Student Health

Students sometimes experience stress from academic expectations and/or personal and interpersonal issues that may interfere with their academic performance. If you find yourself facing issues that have the potential to or are already negatively affecting your coursework, you are encouraged to talk with an instructor and/or seek help through University resources available to you.

- U Matter, We Care: If you or someone you know is in distress, please contact umatter@ufl.edu, 352-392-1575, or visit U Matter, We Care website to refer or report a concern and a team member will reach out to the student in distress.
- Counseling and Wellness Center: <u>Visit the Counseling and Wellness Center</u> website or call 352-392-1575 for information on crisis services as well as noncrisis services.
- Student Health Care Center. Call 352-392-1161 for 24/7 information to help you find the care you need, or visit the Student Health Care Center website.
- *University Police Department*: Visit <u>UF Police Department website</u> or call 352-392-1111 (or 9-1-1 for emergencies).
- UF Health Shands Emergency Room / Trauma Center: For immediate medical care call 352-733-0111 or go to the emergency room at 1515 SW Archer Road, Gainesville, FL 32608; Visit the <u>UF Health Emergency Room and Trauma Center</u> website.
- GatorWell Health Promotion Services: For prevention services focused on optimal wellbeing, including Wellness Coaching for Academic Success, visit the <u>GatorWell</u> website or call 352-2734450.

Do not wait until you reach a crisis to come in and talk with us. We have helped many students through stressful situations impacting their academic performance. You are not alone so do not be afraid to ask for assistance.

Policy Related to Guests Attending Class

Only registered students are permitted to attend class. However, we recognize that students who are caretakers may face occasional unexpected challenges creating attendance barriers. Therefore, by exception, a department chair or his or her designee (e.g., instructors) may grant a student permission to bring a guest(s) for a total of two class sessions per semester. This is two sessions total across all courses. No further extensions will be granted. Please note that guests are **not** permitted to attend either cadaver or wet labs. Students are responsible for course material regardless of attendance. For additional information, please review the Classroom Guests of Students policy in its entirety. Link to full policy: https://phhp.ufl.edu/policy-classroom-guests-of-students/

Evaluations

Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at https://gatorevals.aa.ufl.edu/students/. Students will be notified when the evaluation

period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via https://ufl.bluera.com/ufl/. Summaries of course evaluation results are available to students at https://gatorevals.aa.ufl.edu/public-results/.

Class Demeanor

Students are expected to arrive to class on time and behave in a manner that is respectful to the instructor and to fellow students. Chronic tardiness is disrespectful. Please avoid the use of cell phones and restrict eating to outside of the classroom. Opinions held by other students should be respected in discussion, and conversations that do not contribute to the discussion should be held at minimum, if at all.

Inclusive Learning Environment

Public health and health professions are based on the belief in human dignity and on respect for the individual. As we share our personal beliefs inside or outside of the classroom, it is always with the understanding that we value and respect diversity of background, experience, and opinion, where every individual feels valued. We believe in, and promote, openness and tolerance of differences in ethnicity and culture, and we respect differing personal, spiritual, religious and political values. We further believe that celebrating such diversity enriches the quality of the educational experiences we provide our students and enhances our own personal and professional relationships. We embrace The University of Florida's Non-Discrimination Policy, which reads, "The University shall actively promote equal opportunity policies and practices conforming to laws against discrimination. The University is committed to non-discrimination with respect to race, creed, color, religion, age, disability, sex, sexual orientation, gender identity and expression, marital status, national origin, political opinions or affiliations, genetic information and veteran status as protected under the Vietnam Era Veterans' Readjustment Assistance Act."

Course Schedule

	I=				
Week 1	Thurs: Let's Get R-eady to R-umble!				
Aug 22	Reading				
	Wickham and Grolemund Chapter 1.				
	Healy, Preface (follow installation instructions)				
	I refer to the <i>R for Data Science</i> book by the authors'				
	initials WG and the Healy book as Healy .				
Week 2	Tues: Making a Plot				
Aug 27 & 29	Healy Chapter 1, 2, 3				
	WG Chapter 1 (Data visualization)				
	Thurs: Workshop				
Week 3	Tues: R Programming Basics				
Sept 3 & 5	Reading				
	WG Chapter 2 (Workflow: basics)				
	WG Chapter 3 (Data transformation)				
	WG Chapter 4 (Workflow: code style)				
	Thurs: Workshop				
Week 4	Tues: Data Structures and Data Import				
Sept 10 & 12	Reading				
	WG Chapter 5 (Data Tidying)				
	WG Chapter 6 (Workflow: scripts and projects)				
	WG Chapter 7 (Data import)				
	Thurs: Relational Databases				
	Reading				
	WG Chapter 13 (Joins)				
Week 5	Tues: Voter Registration Data				
Sept 17 & 19	Reading				
	Gimpel, Dyck, and Shaw. 2007. "Election-Year				
	Stimuli and the Timing of Registration." Party				
	<i>Politics</i> 13(3): 351-74. (On Canvass)				
	 https://www.brennancenter.org/blog/voter-purge- 				
	rates-remain-high-analysis-finds				
	Thurs: Little Errors and Big Data				
	<u>Reading</u>				
	Enrijeta Shino, Michael Martinez, Michael P.				
	McDonald, and Daniel Smith. 2020. "Verifying Voter				
	Registration Records: Part of Special Symposium on				
	Election Sciences." American Politics Research				
	48(6): 677-81.				
	"Texas Audit Proposed by GOP Would Miss Minor				
	But Real Errors." [<u>Link</u>]				

14/ I- C	Turan Carrana Data		
Week 6	Tues: Census Data		
Sept 24 & 26	Reading		
	<u>TidyCensus package documentation</u> (it reads more		
	like a primer on how to use the package)		
	Thurs: Workshop		
Week 7	Tues: R Markdown and YAML		
Oct 1 & 3	Reading		
	https://rmarkdown.rstudio.com/lesson-1.html		
	(Recommend all the lessons)		
	• https://learn-the-		
	web.algonquindesign.ca/topics/markdown-yaml-		
	<u>cheat-sheet/#yaml</u>		
	Thurs: Workshop		
Week 8	Tues: Making Maps		
Oct 8 & 9	Reading		
	Healy Chapter 7		
_	Thurs: Workshop		
Week 9	Tues: Geocoding		
Oct 15 & 17	Reading		
	Brian Amos and Michael P. McDonald. 2020. "A		
	Method to Audit the Assignment of Voters to		
	Districts." Political Analysis 28(3): 356-71.		
	Thurse Markshan		
	Thurs: Workshop		
Week 10	Tues: Parsing Strings		
Week 10 Oct 22 & 24	Tues: Parsing Strings Reading		
	Tues: Parsing Strings Reading Hardcopy: Chapter 11		
	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings		
Oct 22 & 24	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop		
Oct 22 & 24 Week 11	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-		
Oct 22 & 24	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-Difference		
Oct 22 & 24 Week 11	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-Difference Reading		
Oct 22 & 24 Week 11	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-Difference Reading WG: Chapter 7 Exploratory Data Analysis		
Oct 22 & 24 Week 11	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-Difference Reading WG: Chapter 7 Exploratory Data Analysis https://www.mailman.columbia.edu/research/popul		
Oct 22 & 24 Week 11	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-Difference Reading WG: Chapter 7 Exploratory Data Analysis Https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-		
Oct 22 & 24 Week 11	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-Difference Reading WG: Chapter 7 Exploratory Data Analysis https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation		
Oct 22 & 24 Week 11	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-Difference Reading WG: Chapter 7 Exploratory Data Analysis https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation https://www.huffpost.com/entry/early-vote-		
Oct 22 & 24 Week 11	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-Difference Reading WG: Chapter 7 Exploratory Data Analysis https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation https://www.huffpost.com/entry/early-vote-election-eve-p b 12853864		
Oct 22 & 24 Week 11	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-Difference Reading WG: Chapter 7 Exploratory Data Analysis https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation https://www.huffpost.com/entry/early-vote-election-eve-p b 12853864 Thurs: Campus Early Voting		
Oct 22 & 24 Week 11	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-Difference Reading WG: Chapter 7 Exploratory Data Analysis https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation https://www.huffpost.com/entry/early-vote-election-eve-p b 12853864 Thurs: Campus Early Voting Reading		
Oct 22 & 24 Week 11	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-Difference Reading WG: Chapter 7 Exploratory Data Analysis https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation https://www.huffpost.com/entry/early-vote-election-eve-p b 12853864 Thurs: Campus Early Voting Reading Enrijeta Shino and Daniel A. Smith. 2020.		
Oct 22 & 24 Week 11	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-Difference Reading WG: Chapter 7 Exploratory Data Analysis https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation https://www.huffpost.com/entry/early-vote-election-eve-p b 12853864 Thurs: Campus Early Voting Reading Enrijeta Shino and Daniel A. Smith. 2020. "Mobilizing the Youth Vote? Early Voting on College		
Oct 22 & 24 Week 11	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-Difference Reading WG: Chapter 7 Exploratory Data Analysis https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation https://www.huffpost.com/entry/early-vote-election-eve-p b 12853864 Thurs: Campus Early Voting Reading Enrijeta Shino and Daniel A. Smith. 2020. "Mobilizing the Youth Vote? Early Voting on College Campuses." Election Law Journal 19(4): 524-541.		
Oct 22 & 24 Week 11	Tues: Parsing Strings Reading Hardcopy: Chapter 11 Online: 14 Strings Thurs: Workshop Tues: Exploratory Data Analysis & Difference-in-Difference Reading WG: Chapter 7 Exploratory Data Analysis https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation https://www.huffpost.com/entry/early-vote-election-eve-p b 12853864 Thurs: Campus Early Voting Reading Enrijeta Shino and Daniel A. Smith. 2020. "Mobilizing the Youth Vote? Early Voting on College		

Week 12	Tues: Project Status Check
Nov 5 & 7	Thurs: Project Status Check
Week 13	Tues: Ecological Inference
Nov 12 & 14	Reading
	https://rpubs.com/rjb6233/ei
	Thurs: Workshop
Week 14	Tues: Class Presentations
Nov 19 & 21	Thurs: Class Presentations
Week 15	Thanksgiving Break
Nov 26 & 28	
Week 16	Tues: Class Presentations
Dec 3	Thurs: Reading Day
Week 17	Fri: FINAL EXAM @3-5pm
Dec 9	(We will use this class for presentations, if needed)

Campaign Targeting Project

What is the goal of this project?

This project simulates a campaign's voter targeting plan. Students will select a district or other locality (such as a county or town) and devise a targeting plan for mail, phone, and canvassing operations. A statewide analysis is inappropriate for this exercise. Students must propose a district or other locality to work on, which the instructor must approve.

For the purposes of this assignment, the election must be a general election with a Democrat and Republican running for an office. While the techniques discussed here can be applied to primary and non-partisan elections, this adds a complicated layer of complexity that is beyond this assignment.

The mechanics of this project are the essentially the same for a Democrat or Republican candidate. It is these mechanics that you will be graded on. If you have ambitions to work on a partisan campaign it is recommended that you select a candidate from that party.

Who may work on this project?

This campaign project is tailored for students in the University of Florida's campaigning program. All students may choose to do this project. All students, including campaigning program students, are welcome to devise their own projects in consultation with the instructor. A reminder is that any project must analyze election administration data, such as the precinct results and voter file data that are components of this project. All students are to work independently, unless given approval from the instructor.

What am I required to do?

There are two important components to this project. The first step analyzes precinct election returns to select persuasion and Get-Out-The-Vote (GOTV) targets. The second step analyzes voter files to identify individuals to contact within the persuasion and GOTV target precincts.

Step 1: Identify precincts in your district or locality

You will identify all the precincts within your district or locality. Among these precincts you will identify persuasion and GOTV targets. Precinct election results available from state or county election offices are usually a good source to identify the list of precincts in your district or locality. Generally, precinct election results are organized such that the past results for your office of interest will be reported only for the precincts within the district or locality.

You will need precinct election results for other steps in this assignment. County election officials occasionally change precinct boundaries following a redistricting or for administrative reasons. Local election offices usually provide precinct maps which can be compared to district and locality boundaries to verify there have been no changes to precinct boundaries. If changes have occurred for the upcoming election since the last election was held, discuss these changes and how best to manage them with the instructor.

Step 2: Identify persuasion and GOTV precinct targets

Persuasion targets are places where voters show the most *variability* in their support for the two major parties' candidates in general elections. A common misperception is that the precincts with the most variability are the competitive precincts, those closest to even support for the two major party candidates.

Competitive precincts can be, but are not necessarily, the precincts that make the best persuasion targets. The goal is to identify voters who are most likely to be persuadable. In this context, *variability* is measured by the difference in vote shares for the two major party candidates running for different offices and across different election years.

Typically, consultants want to analyze as many recent elections as they can. However, since 2021 was a redistricting year many precincts changed between 2020 and 2022. Thus, it may not be possible to obtain election results for precinct boundaries consistent across the 2020 and 2022 elections. There are spatial merging techniques to mitigate this issue, but that is outside the scope of this assignment.

Example: Mayor of Targetville

The town of Targetville has five precincts. In the 2022 general election *contested* elections were held for Governor, US House, State House, Mayor, and City Council.

Why examine contested elections? The goal is to measure *typical* candidate support. If one party did not run a candidate, then the major-party vote share will be 100% for the winning candidate. This will not provide a typical measure of the true support for the winning party if an opponent ran. Likewise, campaign consultants may exclude elections where one candidate performed poorly due to atypical events, such as a scandal. Think critically if any uncontested election or unusual circumstances may justify excluding an election from your analysis.

In Table 1, I provide the vote shares for one of the two major party candidates in these elections and across the precinct. The vote shares are calculated only for the Democratic and Republican candidates, omitting any minor party candidates. For example, if the Democrat won 42 votes, the Republican won 58 votes, and the minor candidate won 10 votes, the two major party vote share for the Republican candidate would be calculated as 58/(58+42) = 58% (for simplicity's sake I refer to this as "vote share" hereafter, but I really mean two major party vote share).

Minor party candidates that draw a significant number of votes pose a problem for this exercise, and may be one of the unusual circumstances that justifies exclusion of an election from the analysis.

Table 1 has two important calculations necessary to identify persuasion and GOTV targets. For the persuasion target the important calculation is the standard deviation of the vote shares across the offices. For the GOTV calculation it is the average.

		US	State		City		Standard
Precinct	Governor	House	House	Mayor	Council	Average	Deviation
4	34%	42%	21%	36%	28%	32.2%	8.0%
3	62%	45%	60%	63%	57%	57.4%	7.3%
1	45%	55%	55%	57%	54%	53.2%	4.7%
2	81%	82%	73%	82%	81%	79.8%	3.8%
5	15%	16%	12%	16%	15%	14.8%	1.6%

Table 1. Targetville Precinct Statistics

Note: I spent time making this table presentable. You are graded on presentation style.

The goal of the persuasion target is to identify voters who are most likely to change their party candidate preferences, i.e., may be persuaded by a campaign. The standard deviation measures variability across a list of numbers, with larger values indicating more variability.

Table 1 ranks precincts by their standard deviations. Precinct 4 is the most variable precinct with a standard deviation of 8.0%. Precinct 3 is next with 7.3%. In this contrived example, note how the average vote share in precinct 4 is 32.2% while in precinct 3 it is 57.4%. This is meant to illustrate that the precinct closest to a 50% vote share is not necessarily the one with the most persuadable voters.

How far should one go down the persuasion target list? In the fictional Targetville there are only five precincts but most districts and localities there have many precincts. In practice, campaigns do a cost-benefit analysis to determine how far into the rankings they will go. For the purposes of this project, the first 20 precincts on the ranking will suffice for the persuasion target. If your assigned district or locality has few precincts, please discuss with the instructor.

It is also sometimes the case that precincts will have few registered voters. If a precinct has small population sizes, it may be inappropriate to prioritize in your targets.

The GOTV target works in a similar manner but uses the average in place of the standard deviation. The goal of the GOTV target is to identify the candidate's likely strong supporters. These are the people a campaign most wants to encourage to vote.

In Targetville, Precinct 2 with the highest support for the candidate's party, so this will top the GOTV list. The next two are much more competitive. In practice, there will be a greater range of precinct vote shares. As with the persuasion target, select the top 20 average vote share. It is okay if there is overlap between the persuasion and GOTV target lists. If your assigned district or locality has few precincts, again, please discuss with the instructor.

Step 3: Select Voter File Targets

Once targets have been identified, a campaign contacts potential voters with persuasion and GOTV messages. Voter registration files provide names, addresses, and in some states even phone numbers or email addresses.

Voter files have a precinct identifier. Select all of the registered voters in your target precincts using this identifier (in R the command is filter).

Now it is time to do some microtargeting, or targeting registered based on their individual characteristics.

In a state with party registration you can assume registered Democrats and Republicans will be strong supporters or opponents of your candidate. Those with no party registration (often called "decline to state" or "no party affiliation") are more persuadable. In a state without party registration, consultants often examine past participation in party primaries for clues about a registrant's partisanship.

Past vote history is also important to identify persons likely to vote. A person with a past history of voting in many elections is someone who your campaign will likely target one way or another. You may also wish to consider targeting younger people who are newly registered that haven't had an opportunity to vote in past elections.

Some states, such as Florida, have additional demographic information, such as gender and race or Hispanic ethnicity. Perhaps your campaign has a special reason to target these groups.

In practice, campaign consultants create statistical models to estimate the propensity of a voter to support a partisan candidate and their likelihood of voting. Consultants often supplement individuals' voter registration records with their spending habits from their credit records, their income, more information about their household and neighborhood, any political donations they have made, and so on. The steps outlined here are probably about 80% as effective as a more sophisticated data analytics approach.

Step 4. Create mail merge, phone, or walk sheets.

Programs such as Microsoft Word (the one I am most familiar with) can make a "Mail Merge" document. These documents draw columns from a data file and populate a letter with information found in each row of the data file, such as the name of the person, their address, and so on.

This is your end-product of this assignment. Create a mail merge document ready for printing for your persuasion and GOTV documents. The document must use multiple voter file fields, such as name and address. Any successful mail merge document will satisfy this requirement, such as envelope address labels.

Step 5. Write a memo for the campaign manager

Write a memo for a campaign manager describing the steps you took and assumptions you made in producing the persuasion and GOTV targets. Creating clear and informative tables, as needed, is a plus.

How is my grade for this assignment determined?

You must complete all steps of this assignment to receive a grade. Incomplete projects will receive a failing grade. When submitting your assignment, you will upload all data files and documents you evaluated and produced, including any supporting spreadsheets.

Your grade will be apportioned in the following manner:

Campaign manager memo:

30%

You will be graded on content and professionalism of your memo.

Identifying target precincts:

30%

You will be graded on identifying precincts (Step 1) and identifying persuasion and GOTV targets (Step 2). Document your work and calculations in a spreadsheet.

Identifying target registrants:

30%

You will be graded on the R code and the output it produces.

Produce mail-merge document:

10%

You will be graded on how well the mail merge document achieves its purpose. Any document with a minimum of the registrants' names and addresses is acceptable, such as mailing labels, a personalized letter, or walk or phone sheets. It is not necessary to print this document, but it is necessary to produce a pdf of 100 records (to keep the size of the document manageable). You may use any word processing software.