

POS 6766
3 Credits

Fall 2025 T 8:30am-10:25am
Th 9:35am-10:25am

Informatics
Lecture Room

Election Data Science

Prof. Michael McDonald

Contact Info

Office: Anderson 222

E-mail: michael.mcdonald@ufl.edu

Phone: 352-273-2371

Office Hours

W 9:30am-11:30am or by appointment

Course Description

Introduction to basics of data science including programming for data analytics, file management, relational databases, visualizations, geographic information systems, and web development with application to large-scale election databases.

Course Overview

Campaign consultants since the late 1960s have analyzed large-scale databases in the areas of voting, campaign finance, and redistricting. We learn the properties of election data and the programming tools used to visualize analyses. Students produce a research project published on a public website suitable to present to prospective employers.

Relation to Program Outcomes

This class serves as elective credit for the Political Science Campaign MA program, and the general Political Science MA and PhD degrees. It is most suitable for students with an interest in American politics since the data analyzed is primarily drawn from United States sources. Students with an international interest should speak with the instructor about the suitability of this class to their educational goals.

Student Learning Outcomes

Upon successful completion of this course, graduate students will be able to:

Campaign Field Operation Track (Political Science Campaign MA Program)

1. **Identify** key precinct election results and voter file characteristics to develop strategic campaign objectives.
2. **Analyze** precinct data to distinguish get-out-the-vote and persuasion targets for a legislative district or locality.
3. **Create** targeted voter lists in a format suitable for campaign field operations, demonstrating proficiency in data organization and practical application.
4. **Synthesize** election data into actionable plans, such as canvassing, phone banking, or mass mailing strategies.
5. **Assess** the effectiveness of campaign field operation strategies based on simulated outcomes and contextual variables.

Graduate Research Track (PhD and MA Research-Oriented Students)

1. **Identify** advanced research questions relevant to election administration data and align them with graduate-level research goals.
2. **Assess** the suitability and limitations of large-scale databases, such as voter registration files and precinct-level records, for proposed research.
3. **Apply** R programming to clean, analyze, and interpret complex election administration datasets.
4. **Analyze** large datasets to generate insights and conclusions that contribute to the student's thesis, dissertation, or independent research project.
5. **Synthesize** research findings into professional-grade visualizations presented on an interactive website.
6. **Create** and execute a graduate-level research proposal that demonstrates critical thinking and methodological rigor, with potential for academic publication or professional application.

General Course Outcomes (Applicable to Both Tracks)

1. **Demonstrate** proficiency in R programming through the successful completion of data analysis and visualization tasks based on real-world election examples.
2. **Compare** the unique requirements of campaign field operations and graduate-level research, leveraging course knowledge to address practical and theoretical challenges.
3. **Reflect** on how the analytical and visualization skills acquired in this course align with professional aspirations in politics, data science, or academia.

These outcomes are designed to ensure that students gain both the practical and theoretical expertise necessary for their specialized track while building skills that are transferable across political and data-driven career pathways.

Course Objectives

There are two tracks for graduate students of this course.

The first track is appropriate for students of the Political Science Campaign MA program. Successful students learn the necessary skills to independently conduct a campaign field operation. Students select a legislative district or locality to conduct a campaign simulation. Analyzing precinct election results and voter files, students identify get-out-the-vote and persuasion targets. Students produce lists of targeted voters in a format suitable for canvassing neighborhoods, phone banking, or mass mailing. Detailed instructions for this project are attached at the end of this syllabus.

The second track is appropriate for graduate students that aspire to do graduate-level research. Students must analyze election administrative data, which generally includes large-scale databases such as voter registration files, precincts, or other relevant administrative records. The results must be visualized on a website. In consultation with the professor graduate students make a research proposal and execute the research utilizing the principles taught in the class. The research is at a graduate level, meaning that it is at an advanced level that furthers a student's intended PhD dissertation; or

masters thesis or project. Proposals that further an independent research project for other purposes, such as preparation of a manuscript for academic journal submission, may be considered.

We 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/>. Throughout the course we apply programming tasks to real-world election examples to track R competency.

The analysis and visualization skills you apply here will assist you in your future careers in politics or data science. Alumni of this class have provided their class projects as work product examples, which have helped them in securing a job upon graduation. They have published their research in peer-reviewed academic journals.

Course Materials

- Hadley Wickham and Garrett Golemund, 2nd edition. 2023. [*R for Data Science*](#). Sebastopol, CA: O'Reilly Media Inc.
- Kieran Healy. 2018. [*Data Visualization: A Practical Introduction*](#). 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 Reading Assignment for that day of class.

Additional Academic Resources

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

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 [resources to students in distress](#) 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.

Requirements for class attendance and make-up exams, assignments, and other work in the course are consistent with university policies. See UF Academic Regulations and Policies for more information regarding the University Attendance Policies

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>	<u>% of Overall Grade</u>
Class Project		

Proposal	Sept. 26	10%
Final Submission	Dec. 9	80%
Workshops	Various	10%

Class Assignments Details

Proposal (10% of Overall Grade)

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/question
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 (80% of Overall Grade)

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 do not 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 (10% of Overall Grade)

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	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-
Below 60	E

Your grade can be adjusted further for failure to adhere to the attendance and technology use policies. These grade adjustments are described in detail in the attendance and technology use sections of this syllabus, to reiterate:

- 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.
- Do not 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.

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 [university policies](#).

General University Policies

This course complies with all UF academic policies. For information on those polices and for resources for students, please see [this link](#)

Course Schedule

Week 1 Aug 21	Thurs: Let's Get R-eady to R-umble! <u>Reading</u> <ul style="list-style-type: none"> • Wickham and Grolemund Chapter 1. • Healy, Preface (follow installation instructions) <p>I refer to the <i>R for Data Science</i> book by the authors' initials WG and the Healy book as Healy.</p>
Week 2 Aug 26 & 28	Tues: Making a Plot <ul style="list-style-type: none"> • Healy Chapters 1, 2, 3 • WG Chapter 1 (Data visualization) Thurs: Workshop
Week 3 Sept 2 & 4	Tues: R Programming Basics <u>Reading</u> <ul style="list-style-type: none"> • WG Chapter 2 (Workflow: basics) • WG Chapter 3 (Data transformation) • WG Chapter 4 (Workflow: code style) Thurs: Workshop
Week 4 Sept 9 & 11	Tues: Data Structures and Data Import <u>Reading</u> <ul style="list-style-type: none"> • WG Chapter 5 (Data Tidying) • WG Chapter 6 (Workflow: scripts and projects) • WG Chapter 7 (Data import) Thurs: Relational Databases <u>Reading</u> <ul style="list-style-type: none"> • WG Chapter 13 (Joins)
Week 5 Sept 16 & 18	Tues: Voter Registration Data <u>Reading</u> <ul style="list-style-type: none"> • Gimpel, Dyck, and Shaw. 2007. "Election-Year Stimuli and the Timing of Registration." <i>Party 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> <ul style="list-style-type: none"> • Enrijeta Shino, Michael Martinez, Michael P. McDonald, and Daniel Smith. 2020. "Verifying Voter Registration Records: Part of Special Symposium on Election Sciences." <i>American Politics Research</i> 48(6): 677-81. • "Texas Audit Proposed by GOP Would Miss Minor But Real Errors." [Link]

Week 6 Sept 23 & 25	Tues: Census Data <u>Reading</u> <ul style="list-style-type: none"> • TidyCensus package documentation (it reads more like a primer on how to use the package) Thurs: Workshop
Week 7 Sept 30 & Oct 2	Tues: R Markdown and YAML <u>Reading</u> <ul style="list-style-type: none"> • https://rmarkdown.rstudio.com/lesson-1.html (Recommend all the lessons) • https://learn-the-web.algonquindesign.ca/topics/markdown-yaml-cheat-sheet/#yaml Thurs: Workshop
Week 8 Oct 7 & 9	Tues: Making Maps <u>Reading</u> <ul style="list-style-type: none"> • https://learning.nceas.ucsb.edu/2019-11-RRCourse/spatial-vector-analysis-using-sf.html Thurs: Workshop
Week 9 Oct 14 & 16	Tues: Geocoding <u>Reading</u> <ul style="list-style-type: none"> • Brian Amos and Michael P. McDonald. 2020. "A Method to Audit the Assignment of Voters to Districts." <i>Political Analysis</i> 28(3): 356-71. Thurs: Workshop
Week 10 Oct 21 & 23	Tues: Parsing Strings <u>Reading</u> <ul style="list-style-type: none"> • Hardcopy: Chapter 11 • Online: 14 Strings Thurs: Workshop
Week 11 Oct 28 & 30	Tues: Exploratory Data Analysis & Difference-in-Difference <u>Reading</u> <ul style="list-style-type: none"> • 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 <u>Reading</u> <ul style="list-style-type: none"> • Enrijeta Shino and Daniel A. Smith. 2020. "Mobilizing the Youth Vote? Early Voting on College Campuses." <i>Election Law Journal</i> 19(4): 524-541. (On Canvas)

Week 12 Nov 4 & 6	Tues: Project Status Check Thurs: Project Status Check
Week 13 Nov 11 & 13	Tues: Ecological Inference <u>Reading</u> <ul style="list-style-type: none"> • https://rpubs.com/rjb6233/ei Thurs: Workshop
Week 14 Nov 18 & 20	Tues: Class Presentations Thurs: Class Presentations
Week 15 Nov 25 & 27	Thanksgiving Break
Week 16 Dec 2	Tues: Class Presentations Thurs: Reading Day
Week 17 Dec 8	Fri: FINAL EXAM @3-5pm (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. Calculate votes needed to win

The number of votes needed to win is often the topline number in your memo to a campaign manager or the candidate. This topline statement often takes the form, "Our goal is to win X votes."

To calculate the to-win number estimate the number of people you expect to vote in the upcoming election.

The typical way to do this calculation is to examine *past comparable elections*. Turnout varies across different types of elections, general, primary, presidential, midterm, and local elections, so your best guess is guided by turnout in similar comparable elections.

Once you've identified past comparable elections, obtain the total number of votes for the office your candidate is running for and the total number of registered voters at the time of the election. Divide these two numbers to calculate a voter registration turnout rate. Average this rate across past comparable elections if you have more than one election. Obtain the current registration for the office and multiple by the average registration turnout rate in past comparable elections.

Once you have the total number of registrants expected to vote in the upcoming election, divide by two to calculate your to-win number.

There can be nuances to this calculation. For example, registration counts are not final until the close of the registration period, which may extend through Election Day in same-day registration states. As another example, the past comparable elections may have been uncontested, so you might estimate what would have happened if electoral conditions were similar to the upcoming election.

Step 3: 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.

Precinct	Governor	US House	State House	Mayor	City Council	Average	Standard 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.