

POS 4765	Fall 2025	T 8:30am-10:25am	Informatics
3 Credits		Th 9:35am-10:25am	Lecture Room
Prerequisite:			
Sophomore Standing			
or Higher			

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 major and Data Science major. 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 for the Course

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

1. **Identify** the characteristics and properties of large-scale election administration data, such as voter registration files and precinct-level records.
2. **Assess** the suitability of election administration datasets for specific research questions, avoiding reliance on survey data.
3. **Apply** the R programming language to analyze election administration data and perform data manipulation, cleaning, and summarization.

4. **Analyze** complex election administration data to derive meaningful insights related to real-world electoral processes.
5. **Synthesize** analytical findings into cohesive narratives and visual representations using R and related tools.
6. **Create** an interactive website to showcase the results of a research project, integrating analysis, visualization, and interpretation of election administration data.
7. **Compare** and evaluate the effectiveness of different visualization techniques to communicate election data insights to diverse audiences.
8. **Demonstrate** competency in programming with R through the development of research projects and the completion of practical programming tasks based on real-world election scenarios.
9. **Reflect** on the applicability of learned skills in professional contexts such as politics, data science, graduate studies, or academic publishing.

Course Objectives

Students of this course produce a research project, the results of which are visualized on a website. Students analyze election administration data, which generally includes large-scale databases such as voter registration files, precincts, or other relevant administrative records. It does not include survey data.

Successful completion of this class results in an understanding of the properties of election administration data, how to apply sound analyses to it, and how to visualize the results of analyses.

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 and provided their research in their successful applications to graduate programs.

Course Materials

- Hadley Wickham and Garrett Grolemund, 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)

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)

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.

Class Project (80% of Overall Grade)

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.

The following elements are considered for your project's grade:

- (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.

General University Policies

This course complies with all UF academic policies. For information on those policies 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)