

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

Introduction

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 micro-targeting efforts in 2008. Recent years has 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.

While we focus on election data in this class, 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 program 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.

Readings

- Hadley Wickham and Garrett Grolemund. 2017. *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.

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. Note, the hardcopy and online version of the Wickham textbook does not have the same chapter numbering as the online version. The online version has updated content.

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

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 department 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.

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](#).

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

Grades

Grades consist of a midterm, a final, a paper, a campaign assignment, a redistricting assignment, and classroom participation. The weighting of the aspects of the grade are:

<u>Assignment</u>	<u>Due</u>	<u>% of Overall Grade</u>
Class Project		
Proposal	Sept. 28	10%
Final Submission	Dec. 15	60%
Workshops	Various	30%

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.

The following elements are required:

(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 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. The professor gives students clear direction on visualizations required for their project no later than the Week 13 student meetings.

(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.

A note about the final exam date: The university scheduled different times for the final for the undergraduate (Dec. 13 8-10pm) and graduate (Dec. 15 3-5pm) students. Since there is a class project, I do not anticipate at this time giving a final exam for this class.

For more information on the university's grading policies, see [here](#).

Class Assignments

Proposal

You will write a one to 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 (e.g., 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

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.

You will be graded 70% on content, 20% on style/professionalism of your project, and 10% for your class presentation.

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 election results. These data do not include survey data, and projects analyzing survey data are inappropriate for the class project.

This class is taught concurrently to undergraduates and graduate students. A student's education track determines the appropriate expectations for the project.

All students are expected to meet the requirements outlined above. Graduate students have higher expectations. In addition to the above requirements, graduate students:

- Are expected to select a topic, in consultation with the professor, relevant to their research interests.
- Are expected to have a more thorough literature review as to why their chosen topic is important. This should include references to at least five books or articles outside class readings.
- May conduct original data collection, for example, by contacting election offices.
- 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:

- (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 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 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. There are seven scheduled workshop assignments. You will be allowed to drop one from your 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 guest 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.

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/>.

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 guest 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](#).

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. 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.

Week 1 Aug 24	Thurs: Let's Get R-eady to Rumble! <u>Reading</u> <ul style="list-style-type: none"> • Wickham and Grolemund Chapter 1. <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 29 & 31	Tues: Making a Plot <ul style="list-style-type: none"> • Healy Chapter 2 (Healy Chapter 1 is optional, but an informative read) • Healy Chapter 3 • WG Chapter 3 (Data Visualization) Thurs: Workshop
Week 3 Sept 5 & 7	Tues: R Programming Basics <u>Reading</u> <ul style="list-style-type: none"> • WG Chapter 4 (Workflow: Basics) • WG Chapter 5 (Data Transformation) • WG Chapter 6 (Workflow: scripts) Thurs: Workshop
Week 4 Sept 12 & 14	Tues: Data Structures and Data Import <u>Reading</u> <ul style="list-style-type: none"> • WG Chapter 10 (Tibbles) • WG Chapter 11 (Data Import) Thurs: Relational Databases <u>Reading</u> <ul style="list-style-type: none"> • WG Chapter 13 (Relational Data)
Week 5 Sept 19 & 21	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 26 & 28	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) <p>Thurs: Workshop</p>
Week 7 Oct 3 & 5	<p>Tues: R Markdown and YAML</p> <p>Reading</p> <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 <p>Thurs: Workshop</p>
Week 8 Oct 10 & 12	<p>Tues: Making Maps</p> <p>Reading</p> <ul style="list-style-type: none"> • Healy Chapter 7 <p>Thurs: Workshop</p>
Week 9 Oct 17 & 19	<p>Tues: Geocoding</p> <p>Reading</p> <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. <p>Thurs: Workshop</p>
Week 10 Oct 25 & 27	<p>Tues: Parsing Strings</p> <p>Reading</p> <ul style="list-style-type: none"> • Hardcopy: Chapter 11 • Online: 14 Strings <p>Thurs: Workshop</p>
Week 11 Oct 31 & Nov 2	<p>Tues: Exploratory Data Analysis & Difference-in-Difference</p> <p>Reading</p> <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 <p>Thurs: Campus Early Voting</p> <p>Reading</p> <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 7 & 9	<p>Tues: Ecological Inference</p> <p>Reading</p> <ul style="list-style-type: none"> • https://rpubs.com/rjb6233/ei

	Thurs: Workshop
Week 13 Nov 14 & 16	Tues: Project Status Check Thurs: Project Status Check
Week 14 Nov 21	Tues: Class Presentations Thurs: Thanksgiving
Week 15 Nov 28 & 30	Tues: Class Presentations Thurs: Class Presentations
Week 16 Dec 5	Tues: Class Presentations Thurs: Reading Day
Week 17 Dec 15	Fri: FINAL EXAM @10:00-Noon