DPR 201: Design and Data Analysis for Social Impact

Dr. Miles D. Williams

Spring 2024

E-mail: williamsmd@dension.edu
Office Hours: MW 3:00-4:30 pm
Office: Higley 401C

Web: milesdwilliams15.github.io
Class Hours: MW 1:30-2:50 pm
Class Room: Fellows 203

Course Description

Political data surround us everywhere we look. How do we use this data for maximum social impact? We start by thinking clearly with data and using a good research design. To do these things well, you need to be armed with the appropriate tools and enough conceptual understanding of them to know how to use them.

In this class, we'll cover the basics of research design for social impact. You'll learn about **correlation** and **causation**, how to know when one implies the other, and how to estimate them in the face of **bias** and **noise** that inevitably find their way into our data. Along the way, you'll have opportunities to practice writing about your research.

As long as you apply yourself, by the time you finish this course you will have made progress in your ability to do statistical programming, think clearly with data, and reason quantitatively, all with the goal of applying these skills to problems in the world of politics and policy.

Class Schedule

Notable dates: MLK (22 Jan.), Spring Break (11-15 Mar.), MPSA (04-05 Apr.), Final Day of Instruction (29 Apr.)

Note: BDMF = the Ethan Bueno de Mesquita and Anthony Fowler text. All due dates are by midnight, unless otherwise specified. Course readings for Wednesday discussions will be posted to Canvas.

Getting Started

Week 01, 01/15 - 01/19: Intro to Research Design

- Day 1: Course Intro
- Day 2: Read BDMF Ch1 "Thinking Clearly in a Data-Driven Age" + Getting our tech up and running.

Part I: Correlation

Week 02, 01/22 - 01/26

- Day 1: MLK Day observed [so no class!]
- Day 2: BDMF Ch2 "Correlation: What Is It and What Is It Good for?"

Week 03, 01/29 - 02/02

- Day 1: BDMF Ch4 "Correlation Requires Variation"
- Day 2: Discussion

Week 04, 02/05 - 02/09: Baby Noah Arrived!

Class Postponed

Week 05, 02/12 - 02/16

- Day 1: BDMF Ch5 "Regression for Describing and Forecasting"
- Day 2: Discussion *AC1 due Mon.*

Week 06, 02/19 - 02/23

- Day 1: BDMF Ch6 "Samples, Uncertainty, and Statistical Inference"
- Day 2: Discussion *AC2 due Mon.*

Week 07, 02/26 - 03/01

- Day 1: BDMF Ch7 "Over-Comparing, Under-Reporting"
- Day 2: Discussion *AC3 due Mon.*

Week 08, 03/04 - 03/08

- Day 1: BDMF Ch8 "Reversion to the Mean"
- Day 2: Discussion *AC4 due Mon. + MA1 due Fri.*

Week 09, 03/11 - 03/15

Spring Break!

Part II: Causation

Week 10, 03/18 - 03/22

- Day 1: BDMF Ch3 "Causation: What Is It and What Is It Good for?"
- Day 2: BDMF Ch9 "Why Correlation Doesn't Imply Causation"

Week 11, 03/25 - 03/29

- Day 1: BDMF Ch11 "Randomized Experiments"
- Day 2: Discussion *AC5 due Mon.*

Week 12, 04/01 - 04/05

- Day 1: BDMF Ch10 "Controlling for Confounders"
- Day 2: Discussion *AC6 due Mon.*

Week 13, 04/08 - 04/12

- Day 1: No Class (Eclipse)
- Day 2: BDMF Ch12 "Regression Discontinuity Designs" *AC7 due Mon.*

Week 14, 04/15 - 04/19

- Day 1: BDMF Ch13 "Difference-in-Differences Designs"
- Day 2: Discussion *AC8 due Mon.*

Week 15, 04/22 - 04/26

- Day 1: MA2 Workshop
- Day 2: MA2 Workshop

Week 16, 04/29 - 05/03

- Day 1: Course debrief + evals *MA2 due 05.01 by 4:00 pm*

Course Objectives

In this course, you'll develop a number of skills.

Quantitative You will develop your quantitative skills in the treatment of data. You'll learn how data are gathered, assembled into datasets, and most effectively analyzed to describe and draw inferences for maximum impact.

Communication It is not enough to have a well-designed analysis. To have any social impact with your research you need to be able to effectively communicate your research to others. In this course, you will have opportunities to relay your research and findings through written, verbal, and visual media.

Critical Thinking Critical thinking involves "confronting multiple, competing perspectives and adjudicating between them." This is the essence of what is required to organize data, develop a research design, and draw conclusions.

Statistical Programming In addition to learning the intuition behind statistical methods like multiple regression, causal inference, and hypothesis testing, you will gain experience applying these methods using R and RStudio.

Analysis Software

Students will use R and Posit Workbench (formerly RStudio), which are free and open source. We will use the Posit Workbench (RStudio) version available through a web browser installed on a Denison server: `r.denison.edu`. This can be accessed on campus on any computer or through a VPN off campus. To use the VPN:

1. Go to myDenison.
2. Head to MyApps.
3. Select Remote Access.
4. Follow the instructions to download and use the VPN.

All necessary R packages for the course are already installed in the server version. While using the **Denison server version is highly recommended**, you are also welcome to download and install local versions of R and RStudio on your laptop. Follow this guide.

Required Readings

In addition to reading some short pieces written by journalists, academics, and others available online, we will primarily draw on Ethan Bueno de Misquita and Anthony Fowler's excellent book on quantitative analysis:

> Bueno de Mequita, Ethan and Anthony Fowler. 2021. *Thinking Clearly with Data: A Guide to Quantitative Reasoning and Analysis.* Princeton: Princeton University Press.

I've also compiled an "R Companion" for *Thinking Clearly with Data* that you can refer to as we progress through the course. These follow the main text but, most importantly, provide applied examples in the R programming language.

> An R Companion to *Thinking Clearly with Data*.

Resources

The Instructor (that's me)

During my office hours, my door is always open! If you have any questions or concerns about the course, just drop by my office Monday or Wednesday between 3:00-4:30 pm. No appointment necessary. If those hours don't work for you, we can work out a different time to meet.

You can also email me at williamsmd@denison.edu any time. I'll try to respond as quickly as possible, but if you email me after 5 pm, you may not get a response from me until the next workday. By the end of January, I should have a new small human in my home to look after, too, so you may experience a 24 turnaround on emails sometimes.

Teaching Assistants

DPR hires teaching assistants each semester to provide you with extra help and tutoring. Contact information, hours, and locations for our TAs are included below. Our TAs are a great resource if you're having trouble with your code or are stuck on a problem.

Brennan Kelley

- **email**: kelley_b1@denison.edu
- **hours**: Thursdays 3:00-4:30 pm; Fridays, 2:30-4:30 pm
- **location**: Tables overlooking Atrium in the library

Sury Agrawal

- **email**: agrawa_s1@denison.edu
- **hours**: Mondays 1:30-4:30 pm; Wednesdays 1:30-4:30 pm
- **location**: Olin 216

Accessibility

Students with a documented disability should complete a Semester Request for Accommodations through their Accommodate MyAccommodations app on MyDenison. It is the student's responsibility to contact me privately as soon as possible to discuss specific needs and make arrangements well in advance of an evaluation. I rely on the Academic Resource Center (ARC) located in 020 Higley Hall, to verify the need for reasonable accommodations based on the documentation on file in that office. Reasonable accommodations cannot be applied retroactively and therefore ideally should be enacted early in the semester as they are not automatically carried forward from a previous term and must be requested every semester.

Writing Center

Staffed by student Writing Consultants, the Writing Center is a free resource available to all Denison students. Writing Consultants from a range of majors work with writers one-on-one in all phases of the writing process, including (but not limited to): deciphering assignments, discussing ideas, developing an argument, integrating research and sources, working with faculty feedback, and/or polishing a draft. In addition, Consultants are happy to help with all types of writing, from lab reports, research papers, and informal writing assignments to cover letters, personal statements, and other application materials. The Center welcomes writers from all

backgrounds and levels of college preparation, including multilingual writers. Should a multilingual writer need writing assistance that exceeds the abilities of consultants, the writer can be referred to the Coordinator for Multilingual Learning, Kaly Thayer (thayerk@denison.edu). Writing Center consultations will take place in person in in the Atrium level of the Library; please visit the Writing Center's page (https://my.denison.edu/campus-resources/writing-center) on MyDenison for specific information regarding hours of availability and how to schedule an appointment. The Writing Center strongly recommends signing up for appointments in advance.

Multilingual Support (L2)

Students who use English in addition to other languages are welcome to use the resources available at the Multilingual Learning Office. Kaly Thayer, the Assistant Director of Multilingual Learning, and Anna Adams, the English Language Support Specialist, as well the student consultants who work with them, are trained and experienced in helping students address the different issues that arise when working in more than one language. If English is not your first or only language, please consider utilizing this resource, which is available to ALL Denison students. Ms. Thayer, Ms. Adams, and the student consultants offer a variety of support for L2 students, including consulting with you about your written language (grammar, syntax, word-choices), strategies to manage your reading assignments, assistance with class conversation and presentations, and help devising ways to develop and effectively use all your skills in English. You can set up an appointment via MyDenison - Campus Resources - Multilingual Learning, or by emailing the Multilingual Learning Office directly at englishhelp@denison.edu.

Reporting Sexual Assault

Essays, journals, and other coursework submitted for this class are generally considered confidential pursuant to the University's student record policies. However, students should be aware that University employees are required by University policy to report allegations of discrimination based on sex, gender, gender identity, gender expression, sexual orientation or pregnancy to the Title IX Coordinator or a Deputy Title IX Coordinator. This includes reporting all incidents of sexual misconduct, sexual assault and suspected abuse/neglect of a minor. Further, employees are to report these incidents that occur on campus and/or that involve students at Denison University whenever the employee becomes aware of a possible incident in the course of their employment, including via coursework or advising conversations. There are others on campus to whom you may speak in confidence, including clergy and medical staff and counselors at the Wellness Center. More information on Title IX and the University's Policy prohibiting sex discrimination, including sexual harassment, sexual misconduct, stalking and retaliation, including support resources, how to report, and prevention and education efforts, can be found at: https://denison.edu/campus/title-ix .

R Resources & Cheat Sheets

How to Google R Stuff ◊ `{tidyr}` ◊ `{dplyr}` ◊ `{ggplot2}` ◊ `{rmarkdown}` ◊ `{estimatr}` ◊ `{seerrr}` ◊ `{coolorrr}`

Course Policy

The course policy and requirements are detailed below. It all basically boils down to: (1) show up to class, (2) learn some stuff, and (3) don't cheat or trick me into believing you've accomplished 2.

Grading Policy

Grades at Denison are based on a standard 4.0 scale. You can read more about Denison's grading system here. Generally, a 90 corresponds to an A−, an 80 to a B−, etc.

Grading Scale		
A+: 98%+	A: 92%	A-: 90%
B+: 88%	B: 82%	B-: 80%
C+: 78%	C: 72%	C-: 70%
D+: 68%	D: 62%	D-: 60%
F: below 60		

Attendance and Participation **10%**

You should show up to class and participate! Missing multiple days of class can leave gaps in your learning that make doing assignments down the road more difficult. As an incentive to show up (both physically and mentally) to class, 10 percent of your grade will come from attendance and participation. Your first unexcused absence will be a freebie. After that, you'll lose 1/2 of your attendance and participation grade per each of two subsequent unexcused absences. That means after 3 unexcused absences you'll lose your entire grade for attendance and participation. Also, if you just come to class, but don't engage, you'll lose 1/2 of your attendance and participation grade as well. [**Note**: *Participation does not only look like talking in class. If you don't feel comfortable or confident (but I hope you do!) speaking up, we'll do some work in groups as well. If you participate in your group, that counts as participation, too.*]

Main Assignments **30% Each**

You'll complete two main research projects this semester, each centered on a key set of skills we've covered in class. The first will deal with ***descriptive inference*** and ***prediction***, and the second will deal with ***causation***. The last of these will serve as our final in the class.

8 Analysis Challenges **3% Each**

These are intended to be a fun and challenging way to practice some data reasoning and data analysis skills. You'll be given data or a data-frame and a prompt asking you to perform some analysis with the data. You'll get the necessary data for these at least 5 days in advance of when they are due. These will be graded using a mastery-based approach. That means that if you complete the challenge successfully, that's awesome! You mastered the skills necessary to do the challenge and you get the full 3pts. But if you complete the challenge and it doesn't quite cut it, that's okay. You can try again (and again) until you nail it.

Reading Responses **6% Total**

This is a low stakes assignment. Most weeks on Wednesday we will read a study or article relevant to the material we covered that week. When we do, you will turn in on Canvas a short set of responses to five questions about the assigned reading. These are:

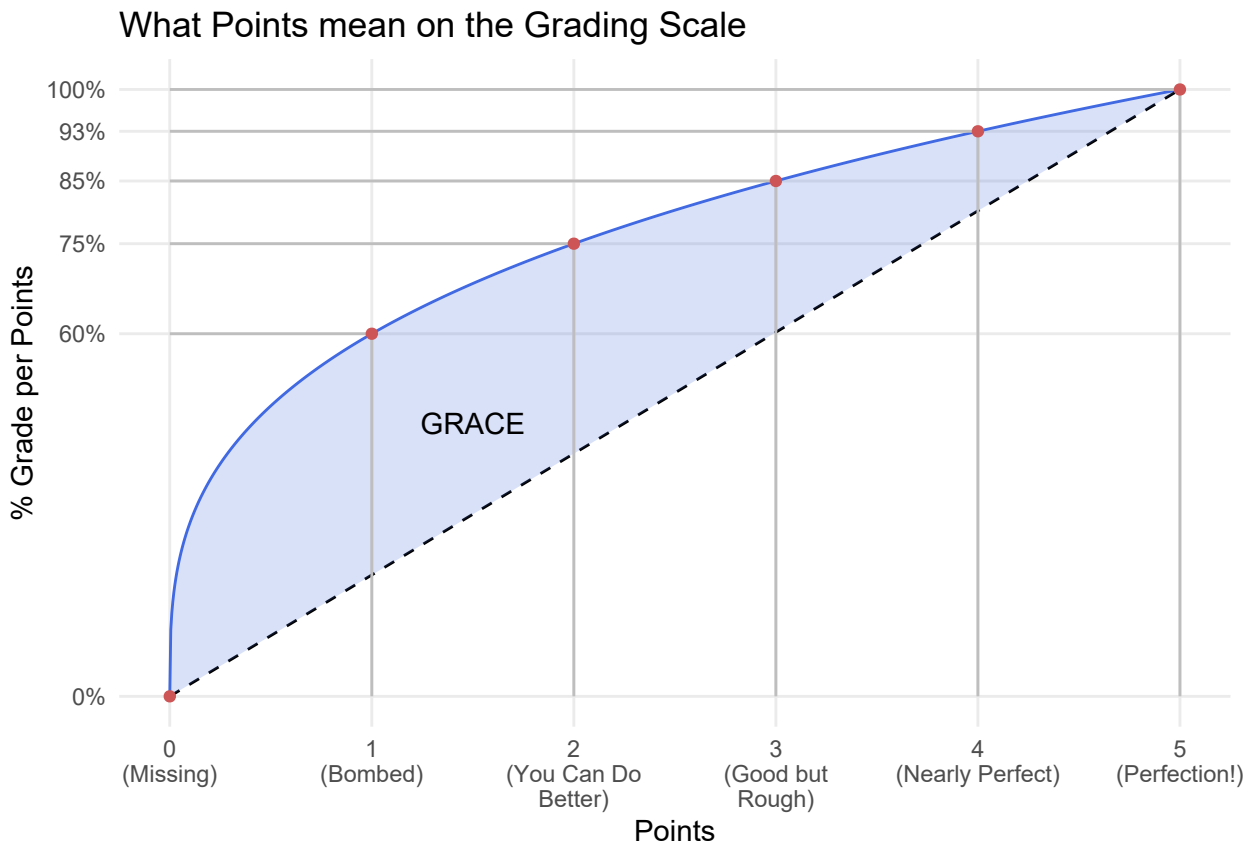
(1) What's the research question?
(2) Why does it matter?
(3) What did the authors do to find an answer? (e.g., what's their research design?)
(4) What did they find?
(5) What are the implications?

Your responses to each question don't need to be longer than 1-2 sentences. They should be sub-

mitted to Canvas Tuesday night before midnight.

$$
\begin{aligned}
6\% &\text{ (Reading Responses)} \\
10\% &\text{ (Attendance and Participation)} \\
60\% &\text{ (Main Assignments)} \\
\underline{+24\%} &\text{ (Analysis Challenges)} \\
100\% &\text{ (Total Assignments)}
\end{aligned}
$$

The following figure provides a summary of my grading strategy and philosophy. For each assignment you submit, I'll evaluate it on a discrete scale from 0 to 5. 5 = "Perfection!" and 0 = "Missing" (e.g., you didn't turn anything in). Your points then get mapped to a percentage point grade for a given assignment according to what I call a "Grace Curve." This means the linkages between my evaluation of your work on the 0-5 scale and your actual grade are not one-to-one. 5 out of 5 = 100% (an "A+"), but 4 out of 5 = 93% (still an "A") and 3 out of 5 = 85% (a "B"), and so on. If you do some simple math, you know that 4 out 5 should actually be 80% and 3 out of 5 should be 60% if I didn't apply the Grace Curve. My grading strategy gives me a simple scale to evaluate your work, but it also ensures that you get some benefit of the doubt. Did you completely bomb an assignment? My instinct is to give you a 1 out of 5, which gives you only 20% for your final grade on said assignment. That seems too harsh. My scale adjusts for my disappointment in the quality of your work and gives you 60% (a "D-") instead. That's still a poor grade, but not one that would be impossible to recover from.



E-mail Policy

I have a simple email policy, and it is targeted at achieving one goal: **maximizing your and my work-life balance**. The policy is this:

> I promise a timely response to **relevant** emails I receive between **9:00 AM** and **5:00 PM** Monday to Friday.

You may not think professors have lives, but we in fact do, and I like to live mine outside of normal working hours. I also will have a new small human to take care of this semester, which makes these hours even more important for me to stick to. That doesn't mean I expect students to abide by chrono-normative standards.[1] But, this does mean that if you email me outside of these windows, I may not respond until the next 9-5 workday.

Make-Up Assigment Policy

There are **NO** make-ups for missed assignments.[2] Don't bother asking. But, if you anticipate having troubles making a due-date and notify me ***in advance***, we can work out a solution.

Second Chances Policy

The analysis challenges have built-in second-chances. That means if your submission didn't show that you clearly mastered the skills necessary to do the challenge well, you can try again. The final project, code glossary, and 4 Qs are a different story. With these, you have one chance. But only one should be necessary. If you come visit me or our TA early and often, you should have no trouble completing these with a passing grade.

Attendance Policy

Attendance is part of your grade (see "Grading Policy"). However, sometimes life happens. There's an emergency, you get sick, you're an athlete and you have to be out of town, etc. Any of these exigencies can be the basis for an excused absence (which means missing class won't count against your final grade). All I ask is that you do me the courtesy of telling me **in advance** of your absence and your justification. After the fact is too late (except in the case of a true and immediate emergency).

Computer-based Excuses

Excuses for late or missed assignments based on CD, flash drive, or hard drive errors are **not acceptable**. The Denison network and server is reliable and accessible. If you use your Google Drive and the RStudio server, all your work will be backed up and easy to access.

Late Assignments

We have a lot of work to do in this class. So turn in your work when it's due. This is meant to help you. I love to procrastinate just as much as anyone else—but if you procrastinate in this course, you will drown. As incentive for keeping on top of your assignments, each day your assignment is late and unexcused (including weekends and holidays) you will lose 5 percent from your final grade for that assignment. There will be no exceptions made for work that is submitted only minutes after a deadline. If something is due by midnight on a Friday (12:00 AM), the moment the clock strikes 12:01 AM your assignment is a day late and you will automatically lose 5% from your grade.

[1] By this, I simply mean classic societal expectations about working vs. leisure hours.

[2] Notwithstanding extenuating circumstances or true emergencies.

The exception to this rule is if a student and I have worked out an arrangement for submitting an assignment at a later date. So, be proactive. If a deadline really won't work for you, I'm more than happy to make accommodations in advance.

Electronic Submission

You will submit all of your assignments electronically via Canvas.

Academic Dishonesty Policy

Don't cheat. Just don't do it.

It should go without saying, but ***plagiarism*** is also a form of cheating and it includes:

1. Copying or paraphrasing the ideas of others without citation or attribution.
2. Copying or paraphrasing the ideas of ***other students in the class***.

I've had to deal with students plagiarizing before. It's painful for me and it puts a blight on the record of the student. It's not only cheating, it's stealing.

When in doubt about whether something constitutes cheating, consult Denison's Code of Academic Integrity.[3] Be advised that this same Code of Academic Integrity requires that instructors notify the Associate Provost of cases of academic dishonesty. **Any incidence of academic dishonesty will result in failure of the course and referral to the Denison judicial process.**

Academic Dishonesty and Generative AI

Remember the previous section where I said you shouldn't cheat? That also applies to GenAI (Generative Artificial Intelligence) tools like ChatGPT. ***But***, there's some nuance to my attitude about GenAI and academic dishonesty. GenAI is a powerful *tool*, and it's a tool that you need to learn to use well because, let's face it, a lot of other people are using it, too. So I want to be very clear that ***I am okay if you use GenAI tools to help you with your work (especially coding in R).*** I've started using ChatGPT to help me with some aspects of my own research and programming. It would be unfair for me to hold you all to a different standard than I hold myself.

So, here's my policy on using GenAI: You can use it, but I want you to be honest about it, and I DO NOT want you to use it uncritically or unthinkingly. What does this look like? **Do not** copy and paste responses or output from tools like ChatGPT and pass them off as your own. **Do** use ChatGPT if you get stuck and need help (just make sure you tell me about it, and that you add your own thoughts, writing, and ideas into the mix as well).

Ultimately, I hope this policy strikes the right balance between honesty and exploration of the potential uses of GenAI while ensuring your own "sweat equity" (effort) goes into your course work as well. I want you to be honest if and when you use GenAI because I'm genuinely curious about its applications. I'm still learning how to use it, too, and I hope that we all can learn how to use it better *together.* Further, I want you to use it critically and thoughtfully because it doesn't work perfectly. While it's powerful (and shockingly so), it is not all-powerful. You need to put in plenty of your own work as well.

You'll probably see a wide variety of attitudes toward GenAI from different faculty on campus. Some attitudes will differ from mine. As instructors, we're still trying to figure out the best policy. Have patience with us!

[3] Of course, if you have to ask yourself if something counts as cheating, then it probably is...