

# DPR 201: Data & Design for Social Impact

Dr. Miles D. Williams

Fall 2022

E-mail: [williamsmd@dension.edu](mailto:williamsmd@dension.edu)

Office Hours: MW 3:00-4:30 pm

Office: Higley 401C

Web: [milesdwilliams15.github.io](https://milesdwilliams15.github.io)

Class Hours: MW 1:30-2:50 pm

Class Room: Higley 029 (in the basement)

---

## 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 which to use and how.

In this class, we'll cover the basics of research design. 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.

## 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 RStudio, which are free and open source! We will primarily use the 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 Mesquita and Anthony Fowler's excellent book on quantitative analysis:

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

## Resources

### We Have an TA!

Will Duquette is the Teaching Assistant for DPR 201. If you have questions about assignments or need help troubleshooting code, Will is an excellent resource.

**Office Hours:** 11:00 am – 1:00 pm

**Location:** Higley 408

**Contact:** [duquet\\_w1@denison.edu](mailto:duquet_w1@denison.edu)

### 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](mailto:thayerk@denison.edu)). Writing Center consultations will take place in person 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](mailto: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}` `{seerrrr}`  
`{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		

## Code Glossary 10pts

We're going to be using statistical programming software in this class, which means you're going to need to learn some new R code. Taking notes and building a repository of code that you can draw on in future work is essential. That's why I want you not only to take notes in class as we write code but also to build a code repository that includes examples of different kinds of data analysis and how to process the data used to do said analysis. The good news is, you won't have to come up with these examples on your own. If you code along with me in class, you'll have a complete code glossary by the end of the semester. This will be due April 23, and I'll check in with you periodically to make sure you're keeping on pace.

**Final Project - Broken down to Four Parts 60pts Total**

Throughout the semester you'll work on a research project. The final product will be due as your "final exam," but you'll actually do the work in segments with different parts due at different points in the semester. Breaking the assignment down this way does two things. First, it ensures that you make steady progress in completing the final assignment. Second, it lets me see how your research project is coming along early and often. The final project will be broken down to four parts. The first three submissions are worth 10pts each and are graded on a pass-fail basis. The goal is to turn something in to me so that I can provide feedback. The final submission will be worth 30pts and will be graded based on a number of criteria, including clarity of writing, having a clear research question, description of your research design, and presentation of your results.

**8 Analysis Challenges 3pts 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. Once you do, you'll get the full 3pts. Until then, you'll get 1pt as a place-holder.

**4 Qs 6pts Total**

This is a low stakes assignment that will be assigned on a rotating basis. Each week on Wednesday we will read some study or article relevant to the material we'll cover that week. In a given week ~3 of you will turn in on Canvas a short set of responses to four questions about the assigned reading. These are:

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

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

$$\begin{array}{r}
 6pts \text{ (4 Qs)} \\
 10pts \text{ (Code Glossary)} \\
 60pts \text{ (Final Project)} \\
 + 8 \times 3pts = 24pts \text{ (Analysis Challenges)} \\
 \hline
 100pts \text{ (Total Assignments)}
 \end{array}$$

**Wait! What about attendance?** You're all young adults. I expect you to come to class and participate. But I won't be taking attendance (at least not formally). If you can't make it to class for some

reason, then you can't make it to class. But if you start missing class regularly, your grade will start to suffer. It won't suffer because you're losing "participation" points. It'll suffer because you're missing out on information covered in class that's necessary for completing analysis challenges, keeping up with your final project, and building your code glossary.

## 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). That doesn't mean I expect students to abide by chrono-normative standards.<sup>1</sup> 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 Assignment Policy

There are **NO** make-ups for missed assignments.<sup>2</sup> 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

Come to class. Attendance is not part of your grade, but poor attendance will lead to poor grades.<sup>3</sup>

## 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.

---

<sup>1</sup>By this, I simply mean classic societal expectations about working vs. leisure hours.

<sup>2</sup>Notwithstanding extenuating circumstances or true emergencies.

<sup>3</sup>Do you think this is just correlation, or is it causal?

## 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.

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*.

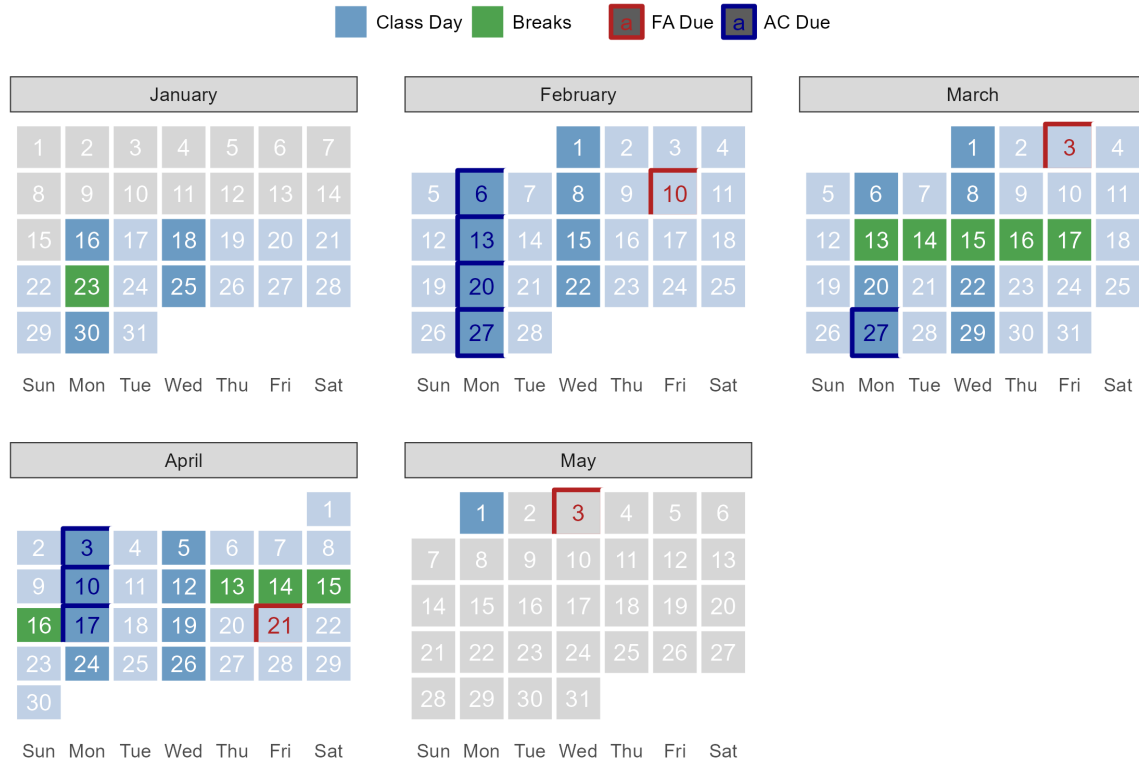
I've had to deal with students plagiarizing before. It's painful for me and it puts a blight on the record of the student. It's not only cheating, it's stealing.

When in doubt about whether something constitutes cheating, consult Denison's [Code of Academic Integrity](#).<sup>4</sup> 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.**

---

<sup>4</sup>Of course, if you have to ask yourself if something counts as cheating, then it probably is...

## Class Schedule



Notable dates: MLK Day Observed (23 Jan.), Spring Break (11-20 Mar.), Final Day of Instruction (01 May)

*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, 01/16 - 01/20: 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. Optional reading on Canvas.

## Part I: Correlation

### Week 02, 01/23 - 01/27

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



**Week 03, 01/30 - 02/03**

- Day 1: BDMF Ch4 “Correlation Requires Variation”
- Day 2: Discussion *AC1 due next Mon.*

**Week 04, 02/06 - 02/10**

- Day 1: BDMF Ch5 “Regression for Describing and Forecasting”
- Day 2: Discussion *Research Q due Fri. and AC2 due Mon.*

**Week 05, 02/13 - 02/17**

- Day 1: BDMF Ch6 “Samples, Uncertainty, and Statistical Inference”
- Day 2: Discussion *AC3 due Mon.*

**Week 06, 02/20 - 02/24**

- Day 1: BDMF Ch7 “Over-Comparing, Under-Reporting”
- Day 2: Discussion *AC4 due Mon.*

**Week 07, 02/27 - 03/03**

- Day 1: BDMF Ch8 “Reversion to the Mean”
- Day 2: Discussion *Intro and Why Care due Fri.*

**Part II: Causation**

**Week 08, 03/06 - 03/10**

- 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 09, 03/13 - 03/17**

*Spring Break!*

**Week 10, 03/20 - 03/24**

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

**Week 11, 03/27 - 03/31**

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

**Week 12, 04/03 - 04/07**

- Day 1: BDMF Ch12 “Regression Discontinuity Designs”
- Day 2: Discussion *AC7 due Mon.*

**Week 13, 04/10 - 04/14**

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

## Part III: Finalizing Projects

### Week 14, 04/17 - 04/21

- Day 1: Work Day
- Day 2: Work Day

*Data and Design due Fri.*

### Week 15, 04/24 - 04/28

- Day 1: Work Day
- Day 2: Work Day + *Course Eval*

### Week 16, 05/01 - 05/05

- Day 1: *Presentations*

*Final research projects due 05.03 by 4:00 pm*