

PAI 741: Data Wrangling and Visualization

Course Syllabus

Prof. Jack Reilly

Fall 2026

Course Information

Class

Time: Tuesdays & Thursdays, 11 AM-12:20 PM

Location: Eggers 225B

Instructor

Professor Jack Reilly

Office: Eggers 225

Office Hours: Friday, 11 AM - Noon and 3-4 PM

Appointments: [schedule online](#)

Phone: 315-443-2687 (office)

email: jreilly@syr.edu

Course Description

Every step in policymaking relies on data. This course introduces students to data management, wrangling, communication, and visualization in the context of public policy, public administration, and behavioral science, as well as the technical tools necessary to do such work in an open and reproducible fashion.

Expanded Description

Data preprocessing, wrangling, and management often consume a large fraction of the time spent doing quantitative data analysis in public administration, public policy, and behavioral science research. Yet these topics frequently do not receive regular attention in methodological courses that focus on statistical inference. This class introduces students to the technical tools necessary to do these tasks in an open and reproducible fashion suitable for modern computational data workflows in the public sector. Throughout the

semester, students will learn the principles and practice of conducting reproducible quantitative research, including readable programming and coding, version control, methods of documentation, data storage, workflow management, and exploratory data visualization. A variety of relevant open-source technical tools will be introduced and used, including but not limited to R (and RStudio), Git (and GitHub), Markdown, and a variety of helper programs to tie things together. Special attention will be paid to data frequently used in public policy, public administration, and behavioral science.

Prerequisites

No formal prerequisites. It is assumed that you have either previously taken or are currently enrolled in an “Introduction to Statistics” or “Quantitative Methods” class at the graduate or undergraduate level (i.e., PAI 721 or MAX 201), and are conversant enough in statistics to be able to work with concepts like “mean” and “standard deviation”.

While this course has no *formal* prerequisites, it does have a substantial *informal* prerequisite: **motivation**. Learning a programming language is challenging work, and students must be prepared to invest the appropriate time, energy, effort, and - *above all* - patience.

Learning Objectives

1. Demonstrate capability in open science and contemporary reproducible data processing, wrangling, and management tools
2. Apply appropriate principles of data and file management to data projects
3. Create effective, reproducible, and well-designed data visualizations with appropriate tools
4. Analyze large-N datasets commonly used in public policy and behavioral science

Materials

Books

- Required:
 - Weidmann, Nils. *Data Management for Social Scientists*. Open access: <https://doi.org/10.1017/9781108990424>
 - Healy, Kieran. *Data Visualization: A Practical Introduction*. Open access: <https://socviz.co>
 - Hadley Wickham, Garrett Grolemund, and Mine Çetinkaya-Rundel, *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*, 2nd ed. Open access: <https://r4ds.hadley.nz>.
- Optional:
 - *Recommended*: Braun & Murdoch, *A First Course in Statistical Programming*, 3rd Edition. Purchase links: [cambridge](#) [amazon](#)
 - Freeman & Ross. *Programming Skills for Data Science*

A Programming Book

There are, in essence, three kinds of books that are useful for the class: a book on data management, a book on data visualization, and a book on data in R: Weidmann's **DMSS**, Healy's **DV**, and Wickham et al's **RDS** respectively.

You may also wish to have a book that goes deeper into the programming logic of R itself. I recommend Braun and Murdoch's (**FCSP**), which is a good general overview of the R language from a statistical programming perspective. Freeman & Ross (**PSDS**) is a more general introduction to the overall data science technical environment, and while it is getting a little dated (2018), it still has useful material on the overall open science research pipeline and associated tools.

Computing

You will need access to a personal computer for this class. It will need to run a full operating system on which you can install local applications outside of app stores and have access to the command line. Windows, macOS, and Linux are all fine. Tablet or web-book OSes - like Chromebooks or iPads - won't be sufficient. Aside from the computer, all significant software we use will be free/open-source, and we'll cover usage and installation in class.

Online Course Resources

Blackboard is our internet-based course platform: <http://blackboard.syr.edu>. In it, you will find submission portals for assignments and a link to the [course webpage](#), where you can find the course syllabus, problem sets, and links to readings.

You can also link to our [course drive here](#), which contains lecture slides, data sets, and some other useful things for the class.

Please note that class itself is the primary source of course-related announcements and material.

Requirements

Overview

Satisfactory completion of the course requires completion of the following:

- Regular course attendance and participation (10%)
- Weekly Assignments (20%)
- Core Exam (15%)
- Practicum 1 (15%)
- Practicum 2 (15%)
- Final Project (25%)

Early assignments will be submitted via Blackboard; once we learn GitHub, later assignments will be submitted through GitHub.

Attendance

One of the guiding principles of my class is that you are adults, and thus, capable of managing your own time. I have little interest in policing your lives. Attendance is taken for each day of class, but you will lose no points on attendance if you happen to miss a day or two: everyone has things that occasionally come up in life that need to be dealt with, and I fully realize that some of those things are things you - very understandably - may not want to discuss with your professor. *That's OK!*

That said, attendance in class is an important element to doing well in the course. If you must miss more than a couple of days, it's a good idea to check in with me so that I don't treat the absences as chronic.¹ The easiest way to do this is just email me with a brief reason when something comes up and you have to miss class (which will also allow me to tell you if you're missing anything particularly important).

If you must miss class, the way to make up what you've missed is straightforward: make sure to look over the posted material, do the reading, get notes from a friend, and still complete the assignment if you are able (and make sure to look over any assignment solutions). If you do these things and still feel like you're missing something, please feel free to come to my office hours and we can talk it through.

Participation

There is no formal grade for "participation". However, I reserve the right to dock a couple points here if you do ridiculous/unprofessional things in class (always coming in late, regularly distracting others, spontaneously breaking out into ribald song in the middle of class, etc).

Assignments

There is an assignment each week in class, **due Thursday by class time**.

Assignments will vary in nature: some will be one-off problem sets while some may build on problem sets from a prior week. All material needed for an assignment will be covered by the lecture material before the assignment is due (usually much earlier), and the assignment itself will be given a week ahead of time. **No assignment work is accepted after class**, as we will go over answers for assignments in class.

Students may miss up to two assignments with no penalty. Students may also work together on assignments - in fact, I encourage you to do so - although each student is ultimately responsible for their own learning and work.

Effort-Based Evaluation

Assignments are evaluated based upon effort and a check completion system. Students who answer every problem will earn a check, with each check worth one point toward their final assignment grade.

¹Chronic absenteeism, for the purpose of the class, is four or more unexcused absences.

Practicums

A practicum is a large assignment that carries more weight and is graded on a scale.² It is untimed, take-home, cumulative, and will be completed on your own time (and computer). Unlike the weekly assignments, you are not allowed to work together on the practicum. (You can think of it as a take-home test, if you like.)

Core Exam

The core exam is a pen-and-paper test, completed without computation. You can find a [study guide here](#) with more information.

Final Project

A project utilizing data of your own choice. You can find more information on the [final project page](#).

Expectations

Etiquette & Decorum

This is a graduate course: I take it for granted that you have a basic interest in the material, an enthusiastic attitude toward participation, and a respectful attitude to everyone in the room. A university classroom is fundamentally a learning community: be courteous to fellow students and the professor, don't let yourself be distracted by things on the internet in class, ask questions when the spirit moves you, etc.

Office & Consultation Hours, Appointments

I encourage you to chat with me at any point if you have questions about the course. I make a habit of staying after our regularly scheduled class for a bit to field questions from individual students.

You can also schedule a meeting at your convenience through my website (<http://jacklreilly.github.io>). If you are on campus, you are always welcome to stop by my office: if the door is open, come on in! (Don't feel like you're intruding! I'll tell you if it's not a good time.)

Email

Email is the best way to contact me. I'm usually pretty responsive, but as a baseline, I always aim to get back to you in a modified 24-hour fashion: by the end of the business day the day after you email. So if you email me at 2 PM Tuesday, I'll get back to you by 6 PM Wednesday at the latest; if 10 PM Thursday, by 6 PM Friday; if you email me at 3 PM on Friday, by 6 PM Monday, etc.³

²While regular weekly assignments are evaluated on an effort-based, check/no-check system, on the practicum, you'll receive a certain number of points out of the total, like a test.

³Again: usually I'm much faster! But if you don't hear from me by this baseline, feel free to bump a reminder.

If your email requires a long response, expect me to encourage you to schedule an appointment with me so that we can more effectively discuss the matter.

Assessment

Course grades are determined according to the breakdown listed under “Requirements” above.

Grading Scale

Grade	Grade Point/Credit	Class Points
A	4.000	94 – 100
A-	3.667	90 – 93.99
B+	3.333	87 – 89.99
B	3.000	83 – 86.99
B-	2.667	80 – 82.99
C+	2.333	77 – 79.99
C	2.000	73 – 76.99
C-	1.667	70 – 72.99
F	0	69.99 or less

Course Assessment Plan

Learning Objective	Assessment Measure
Demonstrate capability in open science and contemporary reproducible data processing, wrangling, and management tools	Core Exam
Apply appropriate principles of data and file management to data projects	Practicums
Create effective, reproducible, and well-designed data visualizations with appropriate tools	Final Project
Analyze large-N datasets commonly used in public policy and behavioral science	Weekly assignments

Course Schedule

subject to change

Week	Topic	Tools
	Preliminaries	
1	Course Introduction	Scripts; R and RStudio
2	A Guide to Your Computer	Markup Languages; Quarto
3	File Management & Version Control	Filesystems; git; GitHub
4	A Field Guide to Data	Data Formats; surveys, readr; tidy
5	Structural Data Manipulation	dplyr; srvyr
6	Data Visualization I	Grammar of Graphics; ggplot2
7	Data Visualization II	ggplot2
8	Core Exam (Thursday, October 15)	
9	Workflow & Data Retrieval	tidycensus; APIs; database/SQL overview
10	Social Networks & Network Data	iGraph; statnet
11	Cartography	tidyverse mapping; color scales; projection
12	Geographic Data	sf; tigris; mapgl; osm
13	Grab Bag: Statistics, Text, & Working with AI	Loops, t.test(), lm(), glm()
14	Project Presentations	
F	Finals Week (Project Due)	

Advanced Topics (*If we get to them*)

Week	Topic	Tools
	Text Data & Data Scraping	
	Statistical Models	
	Missing Data	
	Web Apps & Visualization	quarto, shiny
	Programming with AI	Claude Code
	Local LLMs	

Logistical Notes

All content may be found linked from the course content pages. Readings and content are posted for the whole semester, so students may look ahead if they wish. Typically, each week has two class days: the first, a lecture, and the second, a workshop. Each workshop day typically has some computing tasks, reference reading or other material, and a data assignment.

What to expect

Each week of class is organized into a learning module. Each learning module covers a topic relevant to data wrangling, management, and visualization, and goes in three phases: lecture, workshop, and assignment.

1. **Lectures:** Each module's material is first presented through lectures given Tuesday. Occasionally, there will be an at-home workshop to complete as well. You should start every week with these. Each week has companion readings and reference materials for those wishing to dive deeper.
2. **Workshops:** Each module's material is then reinforced through a workshop in our Thursday class session.
3. **Assignment:** Each module concludes with an applied assignment (typically, a problem set, but occasionally something else - a practicum or project). The module's assignment is due the week after the module in class, and we will go over solutions together the day it is due.

Workshops (Live Class Sessions)

During workshop sessions, we cover material in the following order:

1. Go over solutions and material for the assignment due that week. Sometimes, this will primarily involve the professor covering solutions, but other times, it will involve students presenting their work on that week, as well. Be prepared to ask questions and talk about your assignments.⁴
2. Conduct hands-on coding sessions for material in the current week's module.
3. Allow for time to begin the assignment for the current week's module, ask the professor questions, etc.

Where to find materials and how to navigate the course

Each learning module can be accessed here, on the class website. We'll use Blackboard for initial assignment submission before transferring to GitHub a few assignments in. All course materials may be found here, with the exception of some that contain proprietary or copyrighted material that must be kept behind a login.

Campus Academic Resources & Policies

The following reflects the approved campus-wide academic policies of Syracuse University. For more information and detail, please consult the full documentation of University policies [here](#).

⁴Note that means the cadence will feel odd at first: the assignment given during Thursday workshop week 1 is due during Thursday workshop week 2, and thus, we'll go over the assignment given during week 1 during week 2 but *after* we've already had the week 2 lecture, etc. It's done this way to make sure you always have a full week to work on the assignment.

Academic Integrity

As a preeminent and inclusive student-focused research institution, Syracuse University places academic integrity at the forefront of learning and considers it a core value and guiding pillar of education. The University's [Academic Integrity Policy](#) provides guidelines for completing academic work with integrity. Students are required to meet both course-specific and University-wide academic integrity expectations, such as crediting your sources, doing your own work, communicating honestly and supporting academic integrity.

Upholding academic integrity includes the protection of faculty intellectual property. Students should not upload, distribute or share an instructor's course materials, including presentations, assignments, exams or other evaluative materials, without permission. Using websites that charge fees or require uploading of course material (e.g., Chegg, Course Hero) to obtain exam solutions or assignments completed by others and presenting them as your own violates academic integrity expectations in this course and may be classified as a Level 3 violation.

All academic integrity expectations that apply to in-person assignments, quizzes and exams also apply online. Students found in violation of the policy are subject to grade sanctions determined by the course instructor and non-grade sanctions determined by the school or college offering the course. Students may not drop or withdraw from courses in which they face a suspected violation. Any established violation in this course may result in course failure regardless of violation level.

Artificial Intelligence

Based on the specific learning outcomes and assignments in this course, artificial intelligence is permitted on the following: course assignments, practicums, and final projects. Artificial intelligence is *not* permitted on in-class exams. See each assignment, quiz, or exam instructions for more information about what artificial intelligence tools are permitted and to what extent, as well as citation requirements. If no instructions are provided for a specific assignment, then no use of any artificial intelligence tool is permitted. Any AI use beyond that which is detailed in course assignments is explicitly prohibited except when documented permission is granted.

Please also see Syracuse University's [general campus AI policies here](#).

I, as the instructor, may use generative AI tools to help develop course materials. When I do, I take full responsibility for reviewing and verifying all AI-generated content to ensure it is accurate, appropriate and aligned with our learning objectives. All final course materials represent my professional judgment about what will best support your learning.

Attendance Policy

Attendance in classes is expected in all courses at Syracuse University. It is a federal requirement that faculty promptly notify the university of students who do not attend or cease to attend any class. Faculty will use Early-Semester Progress Reports and Mid-Semester Progress Reports in Orange SUccess to alert the Registrar and Financial Aid Office on non-attendance. For more information visit:

- [Faculty: Information for Faculty: Non-attendance or Stopped Attending](#)
- [Students: Information for Students: Non-attendance or Stopped Attending](#)

If a student is unable to participate in-person or virtually for an extended period of time (48 hours or more), the student may request an absence notification from their home school/college Dean's Office or through Student Outreach and Support office. Instructors will be notified via the "Absence Notification" flag in Orange SSuccess.

Barnes Center at the Arch (Health, Counseling, etc.) staff will not provide medical excuse notes for students. When Barnes Center staff determine it is medically necessary to remove a student from classes, they will coordinate with Student Outreach and Support case management staff to provide appropriate notification to faculty through Orange Success. For absences lasting less than 48 hours, students are encouraged to discuss academic arrangements directly with their faculty.

Additional information may be found at [Student Outreach and Support: Absence Notifications](#).

Blackboard

This class will use the Blackboard Learning Management System to house the syllabus, course content, links to external course materials, assignments, quizzes, exams, feedback, and grades. Due dates and times in Blackboard are stored in Coordinated Universal Time (UTC) and displayed for each user based on the time zone setting of their computer and data from their internet browser. The system will always display the time zone being used. If an instructor sets a due date of 11pm Eastern time, a student in the Pacific time zone will see a due date of 8pm.

Information about Blackboard is available on [Answers Blackboard](#); alternatively, you can contact Information Technology Services by sending an email to help@syr.edu, calling 315.443.2677, or in-person at the ITS Service Center, located at 1-227 CST in the Life Sciences Complex. Business hours for the Service Center can be found on the [ITS website](#).

Discrimination and Harassment

Syracuse University does not discriminate and prohibits harassment or discrimination related to any protected category including creed, ethnicity, citizenship status, reproductive health decisions, national origin, sex, gender, pregnancy, disability, marital status, age, race, color, veteran status, military status, religion, sexual orientation, domestic violence status, genetic information, gender identity and/or gender expression, shared common ancestry, or any other status protected by applicable law.

Any complaint of discrimination or harassment related to any of these protected bases should be reported to Sheila Johnson-Willis, associate vice president and chief equal opportunity and Title IX officer in the Office of Equal Opportunity, Inclusion and Resolution Services (621 Skytop Road, Suite 1001, Syracuse, NY 13244; equalopp@syr.edu; 315.443.4018). She is responsible for coordinating compliance efforts under various laws including Titles VI, VII, IX and Section 504 of the Rehabilitation Act.

Diversity

It is the intent of this course for students from all diverse backgrounds and perspectives to be well served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that students bring to this class be viewed as a resource, strength, and benefit. It is also critical to present materials and activities that are respectful of diversity: gender, sexuality, disability, age, socioeconomic status, ethnicity, race, and culture. Your suggestions are encouraged and appreciated. Please let your instructor know ways to improve the effectiveness of the course for you personally or for other students or student groups.

Disability

Syracuse University values diversity and inclusion and is committed to a climate of mutual respect and full participation. There may be aspects of the instruction or design of this course that result in barriers to your inclusion and full participation. I invite students to meet with me to discuss strategies and/or accommodations (academic adjustments) that may be essential to your success in collaboration with the Center for Disability Resources (CDR).

To discuss disability-accommodations or register with CDR, visit disabilityresources.syr.edu, or contact the center at 315.443.4498 or disabilityresources@syr.edu for more detailed information. CDR is responsible for coordinating disability-related academic accommodations and will work with you to develop an access plan. Since accommodations may require early planning and generally are not provided retroactively, please contact CDR as soon as possible.

Faith and Religious Accommodations

Syracuse University's Religious Observances Policy recognizes the diversity of faiths represented in the campus community and protects the rights of students, faculty and staff to observe religious holy days according to their traditions. Under the policy, students are given an opportunity to make up any examination, study or work requirements that may be missed due to a religious observance, provided they notify their instructors no later than the academic drop deadline. For any observances occurring before the academic drop deadline, students must notify faculty at least two academic days in advance. Students register their observances using MySlice.

FERPA

The Family Educational Rights and Privacy Act (FERPA) sets forth requirements regarding the privacy of student records. FERPA governs both the access to and release of those records, known as education records, and the information they contain. Under FERPA, faculty have a legal responsibility to protect the confidentiality of student records. For additional information about FERPA and Syracuse University's FERPA policy, see [Compliance with the Family Education Rights and Privacy Act](#) or contact the Office of the Registrar (315.443.2422).

Health and Wellness

Well-being and mental health are significant predictors of academic success. It is critical to take care of yourself physically and emotionally and to effectively navigate stress, anxiety, and depression. Please familiarize yourself with the range of resources the Barnes Center provides <https://ese.syr.edu/bewell/> and seek out support for mental health concerns as needed. Counseling services are available 24/7, 365 days, at 315-443-8000.

ORANGE Alert

ORANGE ALERT, Syracuse University's crisis notification system, uses text messages, phone, and email alerts to provide rapid notification and instructions to members of the University community in the event of a crisis in progress. In the event of an emergency, please use one of the following numbers to reach us:

From any phone: 315.443.2224 From your cell phone: #78 (#SU) Campus landline: 711

For more information on ORANGE ALERT, including how to update your contact information, visit the [DPS website](#).