Syllabus for PLSC 498-001
Visualizing Social Data (Penn State)
Spring 2026 (Regular Session)

Instruction Mode: In Person
Campus/Location: CEDAR Building 134
Meeting Times: *Tuesdays & Thursdays (see LionPATH/Canvas)*
Meeting Dates (session window): 01/12/2026 – 05/01/2026
First/Last Meeting (planned): Tue 01/13/2026 – Thu 04/30/2026
Credits: 3

Instructor: Jared Edgerton
Email: jfe4@psu.edu
Office: Carpenter 513
Office Hours: Wednesdays 10:00-2:00 and by appointment

Enrollment note: If you need assistance enrolling, please email Angela Hill at amr25@psu.edu.

Course Description

This course is designed for undergraduate students majoring in Social Data Analytics and related disciplines. The primary objective is to introduce students to the tools, principles, and practices for visualizing complex social science data. Students will engage with various visualization techniques both static and interactive and apply these methods through hands-on projects. The course emphasizes effective communication of analytical findings through visual narratives, mapping spatial relationships, and addressing common challenges faced when visualizing social data. Students will leave the course prepared to critically assess and clearly convey data-driven insights through compelling visualizations.

Prerequisites. Students are expected to have an understanding of introductory statistics (equivalent to the knowledge they would gain from PL SC 309 or STAT 200 prior to taking this course).

How this course counts. PLSC 498 satisfies the following:

- PLSC B.A. majors — fulfills the political methodology distribution requirement, and can be used toward the requirement for an advanced political science course or the related area course requirement.
- INTPL majors — can be used toward the supporting course requirement.
- PLSC B.S. majors — can be used toward the advanced political science course requirement or the related area course requirement.
- SODA majors — can be used toward the advanced analytics course requirement.

Student Learning Objectives

By the end of the course, students will be able to:

1. Explain and apply core principles of effective data visualization (encoding, scales, axes, color, layout, annotation).

2. Build reproducible visualization workflows in `R` using `ggplot2` and related tools.

3. Select appropriate visualization designs for common social-science data tasks (amounts, distributions, proportions, associations, time series, geospatial).

4. Identify and correct misleading or unclear graphics; justify design decisions with evidence-based reasoning.

5. Communicate analytical findings through visual narratives, including multipanel figures, captions, and context.

6. Visualize uncertainty and limitations appropriately and transparently.

7. Build a basic interactive visualization using `shiny` to support exploratory analysis or communication.

8. Use Git and GitHub to manage code, track revisions, and submit reproducible lab work.

Required Textbooks

- Claus O. Wilke, *Fundamentals of Data Visualization* (O'Reilly).
- Hadley Wickham, Mine Çetinkaya-Rundel, and Garrett Grolemund, *R for Data Science* (2e) (assigned chapters: **Ch. 1, 3, 5, 6, 9** + short instructor supplements as posted).

Note on readings. Weekly readings are paired with the Tuesday lecture and the Thursday lab. Reading assignments are listed in the schedule and on Canvas.

Required Tools

- A laptop brought to **every Thursday lab** (and strongly recommended on Tuesdays).
- `R` + RStudio (or VS Code) and the `tidyverse` (including `ggplot2`).
- Quarto or R Markdown for reproducible submissions.
- **Git + a GitHub account** (used for labs and projects; setup in Week 1 lab).
- Penn State Canvas for announcements, readings, and submissions.

Course Format (Tuesday Lecture + Thursday Studio Lab)

This course uses a consistent weekly rhythm:

- **Tuesdays (Lecture/Discussion):**
  1. **Open-book quiz (5–10 minutes):** a short quiz reviewing the prior week's concepts and readings. **(Begins Week 2.)**
  2. **Lecture + discussion:** principles, common pitfalls, and worked examples.
  3. **Plot Find & Critique (10–15 minutes):** students locate a real-world plot and analyze it using the week's principles.

4. **Technique demo (as needed):** short `R`/`ggplot2` demonstrations to support Thursday's work.

- **Thursdays (In-class Lab/Studio):**
    1. Guided studio work applying the week's concepts to real (often messy) social-science data.
    2. Instructor check-ins + structured peer feedback.
    3. A reproducible submission (code + figure + brief rationale) due after lab (deadline posted on Canvas).

Evaluation (Grading)

Your final grade will be based on:

- **Participation & in-class engagement (10%)**
  Includes discussions, critique activities, and collaboration during labs.
- **Tuesday open-book quizzes (10%)**
  Short quizzes; lowest scores may be dropped (details on Canvas).
- **Weekly studio labs (35%)**
  Reproducible submissions (code + figure + short written justification).
- **Coding midterm exam (15%)**
  An in-class coding exam completed over **two sessions** (Tuesday + Thursday). See schedule.
- **Final project + presentation (30%)**
  A multi-week visualization project using social data, including a basic interactive component and a final presentation.

Letter grades. Standard thresholds: A (93–100), A– (90–92.9), B+ (87–89.9), B (83–86.9), B– (80–82.9), C+ (77–79.9), C (70–76.9), D (60–69.9), F (¡60).

Academic Integrity

Penn State expects students to complete all coursework in accordance with course-specific rules and university policy on academic integrity. Unless explicitly permitted, you must not submit work that is not your own, and you must not assist others in ways that violate course rules. When in doubt, ask before submitting.

Students with Disabilities

Penn State welcomes students with disabilities into the University's educational programs. If you anticipate needing accommodations, contact Student Disability Resources (SDR) early in the semester and provide the instructor with the approved accommodation letter as soon as it is available.

Course Norms

Be professional and collaborative. Visualization is an iterative design practice: you will revise work based on critique, clarity goals, and audience needs. Bring your laptop, keep your work reproducible, and be ready to explain your design choices.

Course Schedule (Spring 2026 — Tuesdays & Thursdays)

Key dates: No class Tue Mar 10 or Thu Mar 12 for spring break, no class Tue Mar 24 or Thu Mar 26 because instructor is traveling. Take home midterm is due Fri Mar 6. Final presentation is Thu Apr 30.

Schedule note: Readings and labs may adjust slightly; Canvas is the authoritative source for due dates.

Week 01: Jan 13 (Tue) & Jan 15 (Thu) — Syllabus Day + Toolchain Setup (Git/GitHub)

- **Tuesday (Syllabus Day):** Course overview; expectations; grading; how labs and submissions work; examples of strong/weak figures; final project overview; Q&A. **No quiz.**
- **Thursday (Lab): Environment + Git/GitHub setup.** Install/check R/RStudio and required packages; set up Quarto/Rmd template; Git install/config; GitHub basics (clone, commit, push); repo workflow for labs.
- **Readings: None assigned.**

Week 02: Jan 20 (Tue) & Jan 22 (Thu) — Mapping Data to Aesthetics (Grammar of Graphics)

- **Tuesday (Lecture):** First open-book quiz; aesthetics, data types, scales as mappings; plot find & critique.
- **Thursday (Lab):** Building layered plots; aesthetics practice; choosing geoms; basic themes; short written rationale.
- **Readings (for Week 02 quiz and lecture):** *Fundamentals of Data Visualization* Ch. 1–2; *R for Data Science* (2e) Ch. 6 (Workflow: Scripts and Projects); *R for Data Science* (2e) Ch. 1 (Data Visualization: *Prerequisites* through *Saving Your Plots*);

Week 03: Jan 27 (Tue) & Jan 29 (Thu) — Coordinate Systems, Axes, and Getting Data Plot-Ready

- **Tuesday (Lecture):** Cartesian vs transformed axes; when logs help/hurt; axis labeling and interpretation.
- **Thursday (Lab):** Coordinate transforms (`coord_*`), log scales, axis formatting; preparing data for plotting (filter/summarize/compute rates).
- **Readings:** *Fundamentals of Data Visualization* Ch. 3; *R for Data Science* (2e) Ch. 9 (Layers): *Coordinate Systems*; *R for Data Science* (2e) Ch. 3 (Data Transformation).

Week 04: Feb 3 (Tue) & Feb 5 (Thu) — Color as Encoding, Highlighting, and a Source of Error

- **Tuesday (Lecture):** Color scales; perception; categorical vs continuous color; common pitfalls; accessibility basics.
- **Thursday (Lab):** Designing color scales; highlighting vs encoding; color-vision deficiency checks; legends and redundancy.
- **Readings:** *Fundamentals of Data Visualization* Ch. 4 and Ch. 19; *R for Data Science* (2e) Ch. 9 (Layers): skim *The Layered Grammar of Graphics* (for how ggplot design choices "compose").

Week 05: Feb 10 (Tue) & Feb 12 (Thu) — Choosing the Right Chart: Amounts and Comparisons + Tidying for Visualization

- **Tuesday (Lecture):** Chart selection; "directory" mindset; proportional ink; bars vs dots; when heatmaps help.
- **Thursday (Lab):** Amounts in practice: bars, grouped/stacked bars, dot plots, heatmaps; emphasis on labeling and ordering.
- **Readings:** *Fundamentals of Data Visualization* Ch. 5 and Ch. 6; *Fundamentals of Data Visualization* Ch. 17 (focus: proportional ink); *R for Data Science* (2e) Ch. 5 (Data Tidying).

Week 06: Feb 17 (Tue) & Feb 19 (Thu) — Distributions I: Histograms, Density, and Overplotting

- **Tuesday (Lecture):** Distribution questions; binning and smoothing; transparency/jitter; common interpretability issues.
- **Thursday (Lab):** Histograms and density plots; grouped comparisons; overplotting fixes (jitter, alpha, 2D bins).
- **Readings:** *Fundamentals of Data Visualization* Ch. 7 and Ch. 18; *R for Data Science* (2e) Ch. 1 (*Visualizing Distributions*); *R for Data Science* (2e) Ch. 9 (Layers): *Statistical Transformations + Position Adjustments*.

Week 07: Feb 24 (Tue) & Feb 26 (Thu) — Distributions II + Multipanel Thinking

- **Tuesday (Lecture):** ECDF intuition; skewed data; Q–Q interpretation; comparing many groups; introduction to small multiples.
- **Thursday (Lab):** ECDFs and Q–Q plots; faceting strategies; multipanel figure basics; clear captions and labeling.
- **Readings:** *Fundamentals of Data Visualization* Ch. 8 and Ch. 9; *Fundamentals of Data Visualization* Ch. 21 (Small Multiples); *R for Data Science* (2e) Ch. 9 (Layers): *Facets*.

Week 08: Mar 3 (Tue) & Mar 5 (Thu) — Coding Midterm Exam (Two Sessions)

- **Midterm format:** Take home coding exam
- **Review:** *Fundamentals of Data Visualization* Ch. 1–9, 17–19, 21; *R for Data Science* (2e) Ch. 1, 3, 5, 6, 9.

Mar 10 (Tue) & Mar 12 (Thu) — *No Class (Spring Break)*

Week 09: Mar 17 (Tue) & Mar 19 (Thu) — Proportions and Nested Proportions

- **Tuesday (Lecture):** Parts-to-whole reasoning; stacked vs side-by-side; when (and when not) to use pies; nested proportions.
- **Thursday (Lab):** Composition tasks with social data; stacked/normalized bars; (optional) treemap/mosaic; design justification.
- **Readings:** *Fundamentals of Data Visualization* Ch. 10 and Ch. 11; *R for Data Science* (2e) Ch. 1 (*Visualizing Relationships*: categorical & categorical; numerical & categorical).

Week 10: Mar 24 (Tue) & Mar 26 (Thu) — *No Class (Instructor traveling)*

Week 11: Mar 31 (Tue) & Apr 2 (Thu) — Associations Among Quantitative Variables *(Expanded)*

- **Tuesday (Lecture):** Scatterplots; paired data; correlation intuition (Pearson vs rank-based); correlograms; avoiding overclaiming; (brief) dimension reduction intuition as "projection + information loss" (no math).
- **Thursday (Lab):** Scatterplots with smoothers; correlation heatmaps; small-multiple scatterplots; interpretation-first workflow; (optional) PCA visualization as a geometric summary.
- **Readings:** *Fundamentals of Data Visualization* Ch. 12; *R for Data Science* (2e) Ch. 1 (*Visualizing Relationships*: two numerical; three+ variables); *R for Data Science* (2e) Ch. 9 (Layers): skim *The Layered Grammar of Graphics* as a synthesis.
- **Milestone (suggested):** Final project data + cleaning plan due.

Week 12: Apr 7 (Tue) & Apr 9 (Thu) — Time Series and Trends *(Expanded)*

- **Tuesday (Lecture):** Multiple time series; annotation; smoothing and functional forms; trend vs noise; detrending basics; (added) seasonality vs trend and the risks of "connecting the dots" when measurement intervals differ.
- **Thursday (Lab):** Time series design patterns; smoothing choices (rolling mean vs LOESS); clear labeling and event annotation; uncertainty ribbons (preview); (optional) decomposition-style visualization (trend/seasonal/residual) for intuition.
- **Readings:** *Fundamentals of Data Visualization* Ch. 13 and Ch. 14.
- **Milestone (suggested):** Final project storyboard (planned figures + audience + narrative) due.

Week 13: Apr 14 (Tue) & Apr 16 (Thu) — Geospatial Data and Mapping Social Phenomena *(Expanded)*

- **Tuesday (Lecture):** Projections; layers; choropleths vs alternatives; cartograms; mapping pitfalls and interpretation; (added) normalization choices (rates vs counts), ecological fallacy reminders, and when maps are the wrong chart.
- **Thursday (Lab):** `sf` workflow; joins; choropleths; projection choices; legend/binning decisions; (optional) intro to interactive mapping *or* a non-choropleth alternative (dot density / point maps) as a stretch.
- **Readings:** *Fundamentals of Data Visualization* Ch. 15.

Week 14: Apr 21 (Tue) & Apr 23 (Thu) — Uncertainty, Publishing-Quality Figure Design *(Expanded)*, *Shiny App*

- **Tuesday (Lecture):** Uncertainty of estimates and fits; frequencies vs probabilities; communicating limitations; "don't go 3D" and other high-impact design rules; (added) figure checklists for publication (captions, sources, units, transformations, and reproducibility expectations).
- **Thursday (Lab):** Uncertainty in practice (intervals/ribbons); redesign lab emphasizing redundancy, captions, and context; (added) peer-critique workflow + "publication pass" (export settings, aspect ratio, text size, color accessibility).

- **Readings:** *Fundamentals of Data Visualization* Ch. 16; *Fundamentals of Data Visualization* Ch. 20–23 and Ch. 26 (selected sections).

Week 15: Apr 28 (Tue) & Apr 30 (Thu) — Final Project Presentations (Two Sessions)

- **Tuesday (Presentations):** In-class coding session to debug the final presentation.
- **Thursday (Presentations):** Final project presentations + wrap-up and reflection.
- **Deliverables:** Final repo + figures + short write-up + (basic) Shiny app (if assigned earlier).

Final Project (Overview)

Goal: Produce a portfolio-quality visualization product using social data, with a clear audience, narrative, and a reproducible workflow in `R`.

Minimum components:

- A clearly stated question and target audience.
- A documented dataset (source, time period, unit of analysis, limitations). Use one of the data sets I provided during the course.
- At least **3** polished figures demonstrating different skills from the course (e.g., distributions, associations, time series, geospatial, uncertainty).
- Reproducible code (Quarto/Rmd or scripts) that generates the figures end-to-end.
- A **basic Shiny app** (or Shiny-enabled document) that lets a user explore at least one of your figures via interactive controls.
- A short write-up (or slide deck) explaining design decisions, tradeoffs, and limitations.