Data Visualization
140.610.79, Summer Inst. Term, 1 Credit

Last updated June 6, 2025

Instructor

Erik Westlund, PhD
Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health
ewestlund@jhu.edu

Course Websites

- **GitHub:** https://github.com/erikwestlund/data-viz-summer-25
- **CoursePlus:**
  https://courseplus.jhu.edu/core/index.cfm/go/course.home/coid/23902/

Course Description

This course introduces students to the foundational principles and practical skills of data visualization, with a focus on enhancing the clarity and impact of statistical communication.

The course emphasizes visualization as a critical tool for statistical reasoning, model interpretation, and scientific communication. Students will learn to design and evaluate visualizations that are accurate, clear, and audience-appropriate. Topics include common pitfalls in data presentation, best practices for visual storytelling, and the role of graphics in conveying complex analytical insights.

Hands-on exercises will focus primarily on R, using ggplot2 and the *tidyverse*, with supplementary coverage of Python, Stata, and tools for interactive dashboards.

Course Learning Objectives

- Conceptualize and implement effective data visualizations using R and ggplot2.
- Select appropriate visualization techniques based on data type, analytical goals, and audience needs.
- Critically evaluate data visualizations for effectiveness and accuracy.
- Communicate complex data insights through clear and impactful storytelling.
- Incorporate reproducibility and transparency into the data visualization process using modern coding practices.

- Address common challenges and pitfalls in data visualization, including choosing incorrect statistical transformations, misleading uses of scale and range, and poor design.
- Integrate visualizations into workflows for statistical modeling and applied data science.

Intended Audience

Students in public health, biostatistics, data science, or related fields.

Prerequisites

- Experience with elementary statistics and linear modeling
- Basic familiarity with R and RStudio.

Course Schedule

Please see the course website for a comprehensive schedule and list of activities.

Course Book

We will use Kieran Healy's *Data Visualization* as a course text. It is available for free online: https://socviz.co/. You can purchase a physical copy on Amazon if you wish.

Methods of Assessment

- Participation: 20%
  Students are expected to attend all sessions via Zoom and engage actively by asking questions and participating in discussions.

- Problem sets: 30%
  There will be three brief problem sets—one assigned each day—designed to reinforce key concepts and techniques.

- Final Project: 50%
  Students will submit a scientific notebook (e.g., Quarto, Rmarkdown, Jupyter) that presents a comprehensive visualization of a dataset, ideally one relevant to their work; full project guidelines are available on the course website

Course Evaluation

Course evaluations help improve the quality of this course as it develops over time. The School of Public Health aims for a 100% course evaluation rate. Information on how to do so will be provided in class

Generative AI and Code Assistants

Using AI tools for visualization is common and will be a core topic in this course. Use of these tools is thus permitted and, in some circumstances, encouraged. <u>It is nevertheless the responsibility of the student to understand the output of these tools and ensure their correctness</u>. Students are strongly encouraged to approach these tools as learning aids and not crutches.

Academic Ethics and Student Conduct Code

Students enrolled in the Bloomberg School of Public Health of The Johns Hopkins University assume an obligation to conduct themselves in a manner appropriate to the University's mission as an institution of higher education. A student is obligated to refrain from acts which he, she, or they know, or under the circumstances has reason to know, impair the academic integrity of the University. Violations of academic integrity include, but are not limited to: cheating; plagiarism; knowingly furnishing false information to any agent of the University for inclusion in the academic record; violation of the rights and welfare of animal or human subjects in research; and misconduct as a member of either School or University committees or recognized groups or organizations.

Students should be familiar with the policies and procedures specified under Policy and Procedure Manual Student-01 (Academic Ethics), available at this web site, under the fourth heading: https://my.publichealth.jhu.edu/Resources/PoliciesProcedures/ppm/Pages/default.aspx (retrieved June 2, 2025).

The faculty, staff and students of the Bloomberg School of Public Health and the Johns Hopkins University have the shared responsibility to conduct themselves in a manner that upholds the law and respects the rights of others. Students enrolled in the School are subject to the Student Conduct Code (detailed in Policy and Procedure Manual Student-06) and assume an obligation to conduct themselves in a manner which upholds the law and respects the rights of others. They are responsible for maintaining the academic integrity of the institution and for preserving an environment conducive to the safe pursuit of the School's educational, research, and professional practice missions.

Student Health and Well-being

The Bloomberg School of Public Health is committed to the holistic well-being of its students and offers a wide spectrum of support as they progress toward becoming confident, capable, and fulfilled leaders in public health.

If you are struggling with anxiety, stress, depression, or other mental health related concerns, please consider contacting Mental Health Services.

- Student support: https://bit.ly/bsphstudentsupport
- Mental Health Services: https://wellbeing.jhu.edu/MentalHealthServices/

If you need more immediate support, or if you are concerned about a friend who might be in crisis, please contact the Behavioral Health Crisis Support Team (BHCST) at 410-516-9355, 24 hours a day, 7 days a week. The BHCST provides mobile crisis response to Baltimore campuses within the Charles Village, Mt. Vernon, and East Baltimore Public Safety footprints, and consultation for Washington, DC.

Students with Disabilities: Accommodations and Accessibility

Student Disability Services (SDS) provides accessible and inclusive educational experiences for students with disabilities by collaborating with campus partners to proactively remove barriers, raise awareness of equitable practices, and foster an appreciation of disability as an area of diversity while utilizing a wide range of collaborative approaches from individualized accommodations to universal design.

To request disability-related accommodations or services at Bloomberg School of Public Health:

1. Complete the SDS online application via AIM (see link below)
2. Submit documentation using the provided link after application submission.
3. Schedule a meeting with Audrey Ndaba (see contact below) to discuss your needs and accommodations.

Resources:

- SDS: https://publichealth.jhu.edu/about/inclusion-diversity-anti-racism-and-equity-idare/student-disability-services
- AIM: https://hunter.accessiblelearning.com/JHU/ApplicationStudent.aspx
- Audrey Ndaba: https://outlook.office365.com/book/SDSBSPH@live.johnshopkins.edu/

Schedule

Please try to finish reading before class, but you can get by without doing so.

Date/Time/Topic	Due	Notes & Exercises
Tuesday, June 10 9:00–10:50 am	Healy, chs. 1–3	- Why visualize? What makes a good visualization? - Intro to tools: R, Git, GitHub, editors, notebooks - Notebook formats: RMarkdown, Quarto, Jupyter (brief overview) - Core charts: bar charts, histograms, scatterplots - Levels of measurement and why they matter - Hands-on: set up notebook, load data, transform, plot, commit Homework: Problem Set 1 due Wednesday before class
Wednesday, June 11 9:00–10:50 am	Healy, chs. 4–6 Problem set 1 due	- The grammar of graphics (as a concept & as software) - Avoiding common ways visualizations mislead - Visualization as layers: data → aesthetics → geoms → stats → labels - Aggregation: built-in vs. manual - Working with facets and fills - Annotation and design principles (color, readability, clarity) - Hands-on: build a layered and well-labeled plot Homework: Problem Set 2 due Thursday before class
Thursday, June 12 9:00–10:50 am	Healy, chs. 7–8 Problem set 2 due	- Choosing what to visualize: causal questions and lazy DAGs - Generative experimentation - Plot types: time series, forest plots, ridgeline plots, maps - Dashboards: brief intro to Shiny and alternatives (demo or scaffolded build) - Hands-on: extend a basic Shiny app or explore an existing dashboard Homework: Problem Set 3 due Friday at noon
Friday, June 13	Problem set 3 due at noon	
Thursday, July 3	Final project due at 11:59 pm	