

Instructor guide: Data visualization unit

This instructor guide is designed as a flexible framework to help you deliver the unit effectively.

- Each lesson is summarized with key takeaways and includes cues for additional, step-by-step directions for activities.
- Feel free to adapt and customize the content in the slide deck to fit your teaching style and your students' needs by navigating to File → Make a Copy to get started.

Overview & introduction

Summary:

This unit on **data visualization** examines how visual representations of data shape scientific interpretation, communication, and rigor. Students learn how choices in plot type, data display, dimensionality reduction, and design can either clarify underlying patterns or introduce misleading conclusions. Learners develop the skills to select, evaluate, and refine visualizations that accurately reflect data structure, variability, and relationships. This unit is useful for early-career researchers and students seeking to produce clear, transparent, and publication-quality figures.

Why use this unit:

- This unit helps students develop a rigorous approach to data visualization, emphasizing how design choices directly influence interpretation, transparency, and scientific credibility.
- Each lesson pairs conceptual understanding with hands-on activities, enabling learners to diagnose common visualization pitfalls and apply best practices to their own research figures.
- By progressing from critique to construction, the unit supports students in building publication-quality figures that accurately represent data structure, variability, and relationships.

Lesson overviews

Lesson 1: Rigorous Visualization for Clarity and Correct

Lesson summary:

This lesson introduces the dual role of data visualization in science as a tool for communication and discovery. It explores how well-designed visualizations clarify complex findings, while poor design choices (like misleading scales, unnecessary complexity, or inaccessible formats) can distort meaning and obscure the scientific message. Through examples, learners will see how visualization choices directly shape interpretation, and how rigorous practices promote clarity, transparency, and accessibility in research communication.

Goal:

Develop a critical understanding of how to design and evaluate visualizations that communicate data accurately, transparently, and accessibly.

Activity overview:

Students analyze flawed neuroscience visualizations by identifying common issues and explaining what's wrong, then propose specific improvements to strengthen each figure.

Link to activity:

<https://monolith-test-1.vercel.app/activities/osu-dv-01>

Step-by-Step activity instructions:

1. **[Instructor] Students are introduced to a flawed neuroscience data visualization.**
 - a. Encourage students to carefully examine the figure before making selections.
 - b. Prompt them to think about what might be misleading, unclear, or inaccessible.
2. The screen displays the first flawed visualization along with a set of common issues.
 - a. Students select all issues they believe apply (e.g., misleading axes, poor labeling, inappropriate plot type).

3. After making selections, students are prompted to reflect on their choices. They respond to two questions:
 - a. Describe the flaws you identified:
 - b. What improvements would you recommend?
4. Click **CONTINUE** to confirm responses.
 - a. Feedback appears highlighting key issues in the visualization and explaining why they matter.
5. Once complete, click **CONTINUE** to advance to the second scenario.
6. Once complete, click **CONTINUE** to advance to the results screen to view peer responses alongside their own.
7. **[Instructor] Lead discussion while on the results screen.**
 - a. Highlight the visualization pitfalls they encountered and any common ones to discuss further.
 - b. Discuss how these issues could impact scientific interpretation or lead to incorrect conclusions.
 - c. Encourage students to connect these insights to their own figures or research practices.

Activity takeaway:

Identifying common pitfalls in data visualization shows how plot choice might cause misinterpretation.

Lesson 2: Data structures and visualizations

Lesson summary:

This lesson explores how the structure of data (its distribution, pairing, and variability) determines how it should be visualized. Choosing a plot incongruent with expectations of data structure can obscure real findings or create the illusion of effects that do not exist. By contrast, when visualization aligns with data structure, the figure accurately communicates uncertainty, spread, and within-subject relationships. Through examples, learners will examine how plot choice interacts with statistical assumptions and how to select figures that reveal, rather than distort, the truth within data.

Goal:

Develop the ability to recognize how data structure influences visualization choice, and to select plots that faithfully represent variability, distribution, and paired relationships in scientific data.

Activity overview:

Students explore how different plot types reveal or obscure data patterns by toggling between visualizations, then select and justify the plot that best represents each dataset.

Link to activity:

<https://monolith-test-1.vercel.app/activities/osu-dv-02>

Step-by-Step activity instructions:

1. **[Instructor] Students are introduced to the first dataset and its research context.**
 - a. Encourage students to read the dataset description carefully before interacting with the plots.
 - b. Prompt them to think about what they expect the data structure to look like (e.g., variability, pairing, distribution).
2. The screen displays the dataset visualized using an initial plot type.
 - a. Students toggle between different plot options (e.g., Bar Graph $\pm$ SE, Box Plot, Box + Raw Data).
 - b. Encourage students to observe how each visualization reveals or obscures different features of the data.
3. After exploring the plot options, students are prompted to make a selection. They respond to the following:
 - a. Based on what these three plots suggest about data structure, which plot choice would you recommend to most accurately display what you were told about the data?
 - b. Why did you make this selection?
4. Click **SUBMIT RESPONSE** to confirm responses.
 - a. Feedback appears explaining which plot best matches the data structure and why.

5. Once complete, click **CONTINUE TO DATASET 2** to advance to the second dataset.
6. Once complete, click **VIEW GROUP RESPONSES & TAKEAWAYS** to advance to the results screen.
7. **[Instructor] Lead discussion while on the results screen.**
 - a. Discuss how different plot types revealed or obscured important data features (e.g., pairing, outlier, etc.).
 - b. Encourage students to reflect on how they choose plots in their own research.

Activity takeaway:

Data structure should help you identify the most appropriate choice, that works with plot choice expectations, and lead people to see relationships correctly.

Lesson 3: Plot types and what they imply

Lesson summary:

This lesson explores how the choice of plot type conveys implicit assumptions about the relationships within data. Bars, lines, and other visual conventions are not neutral. Rather, they suggest independence or dependence between variables and shape how audiences interpret patterns. Even when supplying correct data, selecting a plot type that could misrepresent those relationships can lead to misinterpretation. By examining examples of common and misleading graph types, learners will develop the ability to match visual representation to data structure, study design, and statistical analysis.

Goal:

Recognize how different plot types communicate assumptions about data relationships, and apply that understanding to choose figures that accurately reflect the structure and intent of a study.

Activity overview:

Learners build a visualization by selecting appropriate plot features for a given dataset, using their choices to support a clear interpretation and draft a study conclusion.

Link to activity:

<https://monolith-test-1.vercel.app/activities/osu-dv-03>

Step-by-Step activity instructions:

1. **[Instructor] Students are introduced to a research scenario and dataset.**
 - a. Encourage students to carefully read the study context and consider what relationships the data may contain.
 - b. Prompt them to think about how different plot features might influence interpretation.
2. The screen displays the dataset along with options for building a visualization.
 - a. Students begin constructing a plot by selecting responses to:
 - i. What comparison do you want to emphasize?
 - ii. How should you display the data relationships?
 - iii. What additional elements should you include?
3. As students make selections, the visualization updates to reflect their choices.
4. After completing the plot, students are prompted to interpret the results. They respond to the following:
 - a. Based on your visualization, write a 1-2 sentence conclusion about aging effects on synaptic plasticity:
5. Click **CONTINUE** to confirm responses.
 - a. Feedback appears explaining how well the selected plot features align with the dataset and whether the interpretation is supported.
 - b. Students see an alternative plot to compare their choices. This plot is not necessarily better, but shows how by selecting different plot choices, the meaning that is conveyed changes.
6. **[Instructor] Lead discussion while on the results screen.**
 - a. Encourage students to reflect on how they construct figures in their own research and whether those choices align with their data structure.

Activity takeaway:

Selecting the right plot requires understanding how plot types encode different assumptions about your data structure and relationships.

Lesson 4: Displaying data variability

Lesson summary:

This lesson focuses on how to accurately visualize the variability of individual data points within datasets. Even when two datasets share identical means and variances, their underlying distributions may tell entirely different stories. This includes whether a dataset is normal, skewed, bimodal, or driven by outliers. By comparing plot types and exploring common sources of variability, learners will understand how figures that reveal distribution and uncertainty communicate more rigorously than those that summarize data through averages alone.

Goal:

Develop skills to visualize and interpret variability in scientific data, using figures that display spread, distribution, and individual data points to ensure transparent and interpretable communication.

Activity overview:

By comparing bar graphs and scatter plots, learners uncover hidden variability, identifying patterns that different visualizations can either mask or reveal.

Link to activity:

<https://monolith-test-1.vercel.app/activities/osu-dv-04>

Step-by-Step activity instructions:

1. **[Instructor] Students are introduced to a set of bar graphs representing data from different studies.**
 - a. Encourage students to review the brief study content and examine the figures for what information might be hidden.
2. Students review the figures and are asked to consider what conclusions they might draw based on these visualizations alone.
3. Click **REVEAL INDIVIDUAL DATA POINTS** to advance.

4. The screen reveals scatter plots of the same datasets. Students now see the underlying data distributions that were not visible in the bar graphs.
5. Students are prompted to identify patterns in the scatter plots. They respond by selecting the pattern that applies (with hints available if needed).
6. Click **CONTINUE TO REAL CONSEQUENCES** to confirm responses.
 - a. Feedback appears explaining the key differences between the datasets and the types of variability present.
7. **[Instructor] Lead discussion while on the results screen.**
 - a. Have students reflect on when it is important to show raw data rather than relying on summary plots.

Activity takeaway:

Showing the variability of individual data points in a figure will more accurately represent the data than mean and variance alone.

Lesson 5: Visualizing high-dimensional data

Lesson summary:

This lesson introduces learners to the concept of dimensionality reduction and outlines the decision making process involved with picking a technique that is appropriate for their use case. It also introduces a handful of example techniques and notes common pitfalls & rigor issues associated with dimensionality reduction.

Goal:

Identify problems created by high dimensionality and solved by dimensionality reduction. Define key factors involved in deciding on which dimensionality reduction techniques to use. List risks posed by improper use of dimensionality reduction.

Activity overview:

Working through real-world research scenarios, learners choose dimensionality reduction techniques that best align with the structure of the data and research goals.

Link to activity:

<https://monolith-test-1.vercel.app/activities/osu-dv-05>

Step-by-Step activity instructions:

1. **[Instructor] Students are introduced to a research scenario involving high-dimensional data.**
 - a. Encourage students to read the scenario carefully and identify the goals of the analysis.
 - b. Prompt them to consider what challenges high-dimensional data might introduce.
2. The screen displays the scenario along with multiple dimensionality reduction technique options.
 - a. Students review the options and consider how each aligns with the data and research goal.
 - b. Students are prompted to make a selection. They respond to the following: Which technique would be most appropriate for finding a low-dimensional representation to use in the decoder?
 - c. Feedback appears explaining why the selected method is appropriate or not, including key considerations.
3. Click **NEXT** to confirm the selection, and advance to the second scenario.
4. Once complete, click **NEXT** to advance to the third, fourth, and fifth scenarios.
5. Once complete, click **VIEW RESULTS** to advance to the results screen where students can access curated tutorials and expert resources.
6. **[Instructor] Lead discussion while on the results screen.**
 - a. Highlight risks of inappropriate dimensionality reduction technique selection.

Activity takeaway:

Dimensionality reduction isn't always necessary, but when used, linear methods work best for linear data while nonlinear dimensionality reduction is better suited for clustering.

Lesson 6: Style choices in data displays

Lesson summary:

This lesson explores how style choices determine whether a figure can lucidly convey information and impact accessibility. Visualizations should be readable and inclusive, using effective graphics that avoid taxing working memory, guide attention, and respect familiar conventions. A visually appealing figure that is, nonetheless, inaccessible (e.g., poor color choices, tiny fonts) fails its purpose. Through examples, learners will examine how style choices impact interpretability.

Goal:

Apply best practices for rigorous data display to ensure readability and accessibility, while developing a sense of when style guidelines for data visualization are relevant.

Activity overview:

Students diagnose design flaws in neuroscience figures and select effective visual fixes, then compare their improved versions to the originals to reinforce best practices.

Link to activity:

<https://monolith-test-1.vercel.app/activities/osu-dv-06>

Step-by-Step activity instructions:

1. **[Instructor] Students are introduced to a flawed neuroscience figure with design issues.**
 - a. Encourage students to examine the figure closely and consider aspects that affect readability and accessibility.
2. The screen displays the figure along with a set of potential design issues.
 - a. Students are shown the present design flaws (e.g., Poor Color Contrast, Not Colorblind Accessible, No Shape Differentiation).
 - b. Students are prompted to select the best design improvement for each identified issue.
3. Click **APPLY CORRECTIONS** to see if selections address the issues.
 - a. Feedback appears explaining why certain design choices do or do not correct the design issues.
 - b. Students may revise their selections if needed.

4. Click **NEXT** to confirm selections.
5. A corrected version of the figure is displayed alongside the original.
 - a. Encourage students to compare the two and note specific improvements.
6. Click **CONTINUE TO FIGURE 2** to advance to the next figure.
7. Once complete, click **CONTINUE TO FIGURE 3** to advance to the final figure.
8. **[Instructor] Lead discussion while on the results for the final figure.**
 - a. Invite students to reflect on how they plan to avoid common design pitfalls.

Activity takeaway:

A well-designed figure can make complex data immediately understandable, while poor design choices can obscure your findings or exclude portions of your audience.

Lesson 7: Building publication-quality figures

Lesson summary:

In this capstone lesson, students are invited to synthesize all visualization principles from the unit by constructing a rigorous, transparent, publication-quality figure. They apply their learning on how choices in plot type, variability display, styling, and panel organization shape the clarity and credibility of scientific communication.

Goal:

Apply learning from throughout the unit to successfully reconstruct a publication-quality figure that displays data clearly and accurately, using rigorous principles of data visualization.

Activity overview:

Drawing on skills from across the unit, learners critique and rebuild a flawed plot by making informed choices about visualization, variability, and design to produce a clearer final figure.

Link to activity:

<https://monolith-test-1.vercel.app/activities/osu-dv-07>

Step-by-Step activity instructions:

1. **[Instructor] Students are introduced to a research question, study context, and sample data.**
 - a. Encourage students to read the context carefully and before examining a flawed figure in detail.
2. Click **CONTINUE TO ANALYSIS** to a screen that displays the flawed plot along with a set of potential issues.
 - a. Students select all issues they believe apply; they must select three relevant issues to advance to the next screen.
3. After making selections, click **CONTINUE TO REBUILD**.
4. Students begin reconstructing by selecting choices for:
 - a. Choose the Plot Type
 - b. Choose the Error Bar Type
 - c. Choose Color Scheme
5. As selections are made, the visualization updates to reflect their choices.
 - a. Feedback appears explaining how the chosen elements improve clarity, accuracy, and interpretability.
 - b. Encourage students to evaluate whether their revised figure more clearly communicates the data and research question.
6. Click **SEE COMPARISON** to confirm selections.
 - a. The improved figure is displayed alongside the original.
7. Click **CONTINUE** to advance to the final screen where students may download a Figure Quality Checklist.
8. **[Instructor] Lead discussion while on the results screen.**
 - a. Encourage students to reflect on how they will apply these principles when creating publication-quality figures.

Activity takeaway:

Improve the accuracy and effectiveness of data visualization by selecting plots that fit data structure and analysis; displaying variability appropriately; using clear, accessible styling; and communicating a coherent story across panels.

Note for Instructors:

Each unit is estimated to comprise approximately 3 hours of instructional time (approx. 15–20 minutes per lesson), but variances in discussion length, student needs, experiences with the interactive activities, or instructor customization may yield different unit and lesson durations.

Observations & final notes

Concepts likely to challenge students:

- Students may have emotional reactions in places where the instructional content differs from their prior experience. This is okay! You're helping them to learn how to do more rigorous work.
 - We advise letting disgruntled students express their points of disagreement, then gently encouraging them to consider why the materials might disagree with what they've been taught previously.
 - **Remember:** There's nothing wrong with asking a student to hold on to their grievance to let you conclude the unit.
 - **Think we got it wrong?** We want to improve! Email us at c4r@seas.upenn.edu.
- Differentiating between exploratory and confirmatory data visualization and the pitfalls of treating patterns in data as confirmed results.
- Understanding the relationship between data structure and relevant plot types that convey expectations about data structure to the viewer.
- Navigating the technical details of multidimensional data visualization and dimensionality reduction.

Final Reminder for Instructors:

This guide is intended as a flexible framework for presenting the unit. Instructors are encouraged to adapt the content to best suit their teaching style and the needs of their students. Feel free to expand on any section, incorporate additional examples, or integrate further interactive elements. Remember, this is your presentation: use it as a starting point and customize it to best serve your teaching.