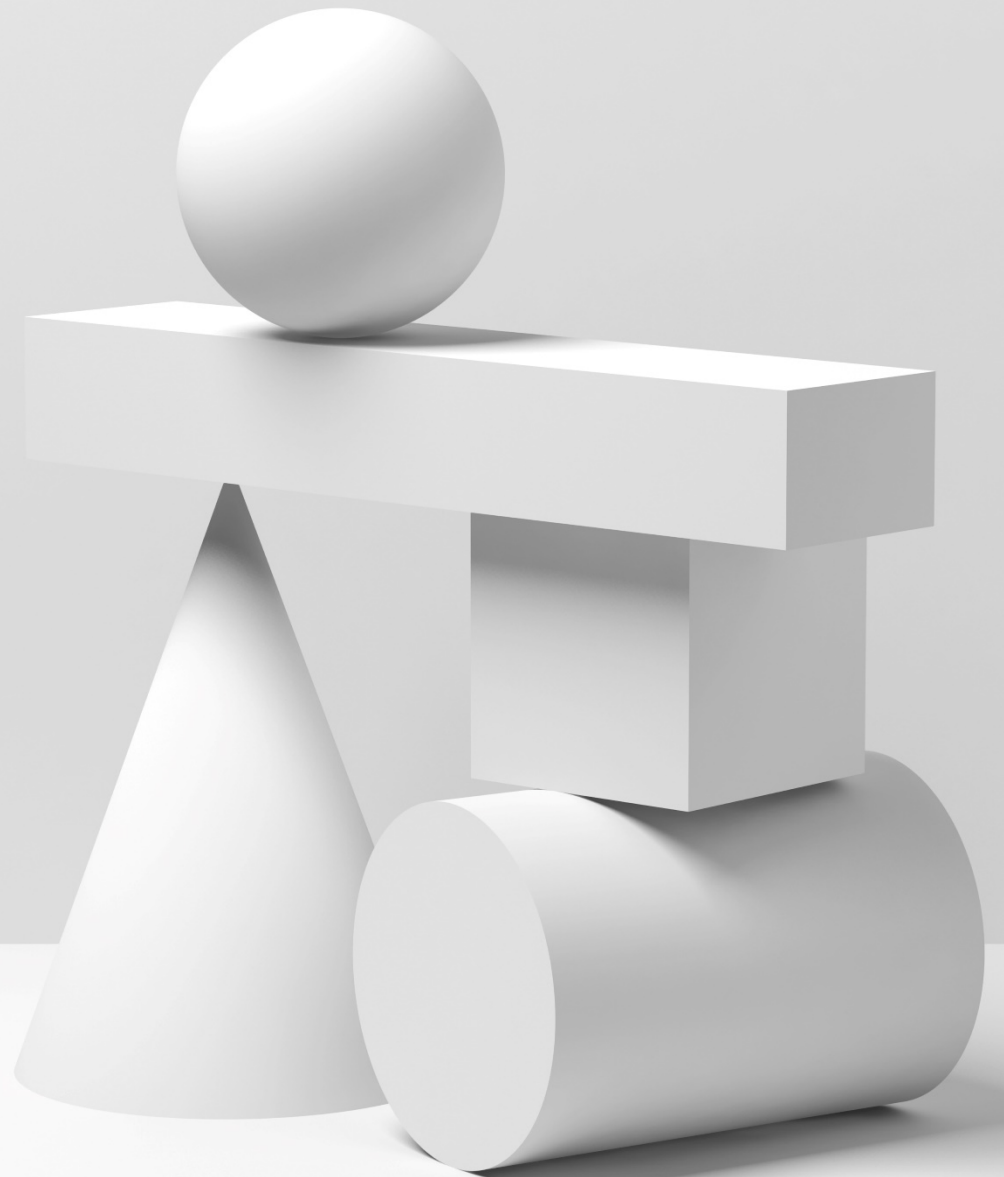


E2A Tool Kit:

VI. Data Visualization

Samia Amin, Carolyn Corea,
Phomdaen Souvanna, and
Siobhan Mills de la Rosa

June 2023





VI. Data Visualization

“Clutter and confusion are not attributes of data – they are shortcomings of design.”
Edward Tufte

A data visualization is a schematic display of numeric data and information. In your work, you have likely encountered numerous examples of bad data visualization: they are hard to interpret and create more questions than answers. While a bad data visualization can create confusion, a high-quality one can build a connection between your audience and the study data. It empowers your audience to further explore the data and apply the evidence in their work. This chapter provides an introduction to planning, developing content for, and designing quality data visualizations. We cover static, animated and interactive data visualization. For all three types, we discuss concepts applicable to making simple ones. Additional resources for developing more advanced data visualization are presented at the end of the chapter.

Key Steps



Step 1. Plan

- 1. Understand your audience’s needs.** At its core, a data visualization’s purpose is to communicate complex data insights clearly and effectively. It is important to understand who you are creating the data visualization for and how they might use it before starting the development process. When thinking about your audience, you should:
 - **Be as specific as possible about who your target audience is and consider what they need to know.** See Chapter II. Planning for E2A Products for more information.

- **Determine your audience’s level of expertise.** Does your audience consist of technical experts (e.g., research staff within agencies) or non-technical professionals (e.g., policymakers or practitioners)? Your audience’s level of technical expertise has implications for your data visualization’s design, level of detail, and format. For instance, researchers will likely be interested in seeing technical details (e.g., confidence intervals). In contrast, policymakers or practitioners will likely be more interested in simplified charts that tell a story and prominently display the key insights and actionable steps.
- **Identify the time constraints of your audience.** In other words, how much time will your audience have or be willing to spend looking at your data visualization?
 - If they are likely to have limited time, you may want to stick to a simple static chart. If they have more time (and interest), an interactive dashboard may be appropriate. We will discuss the difference between static and interactive data visualization in more detail a little later.

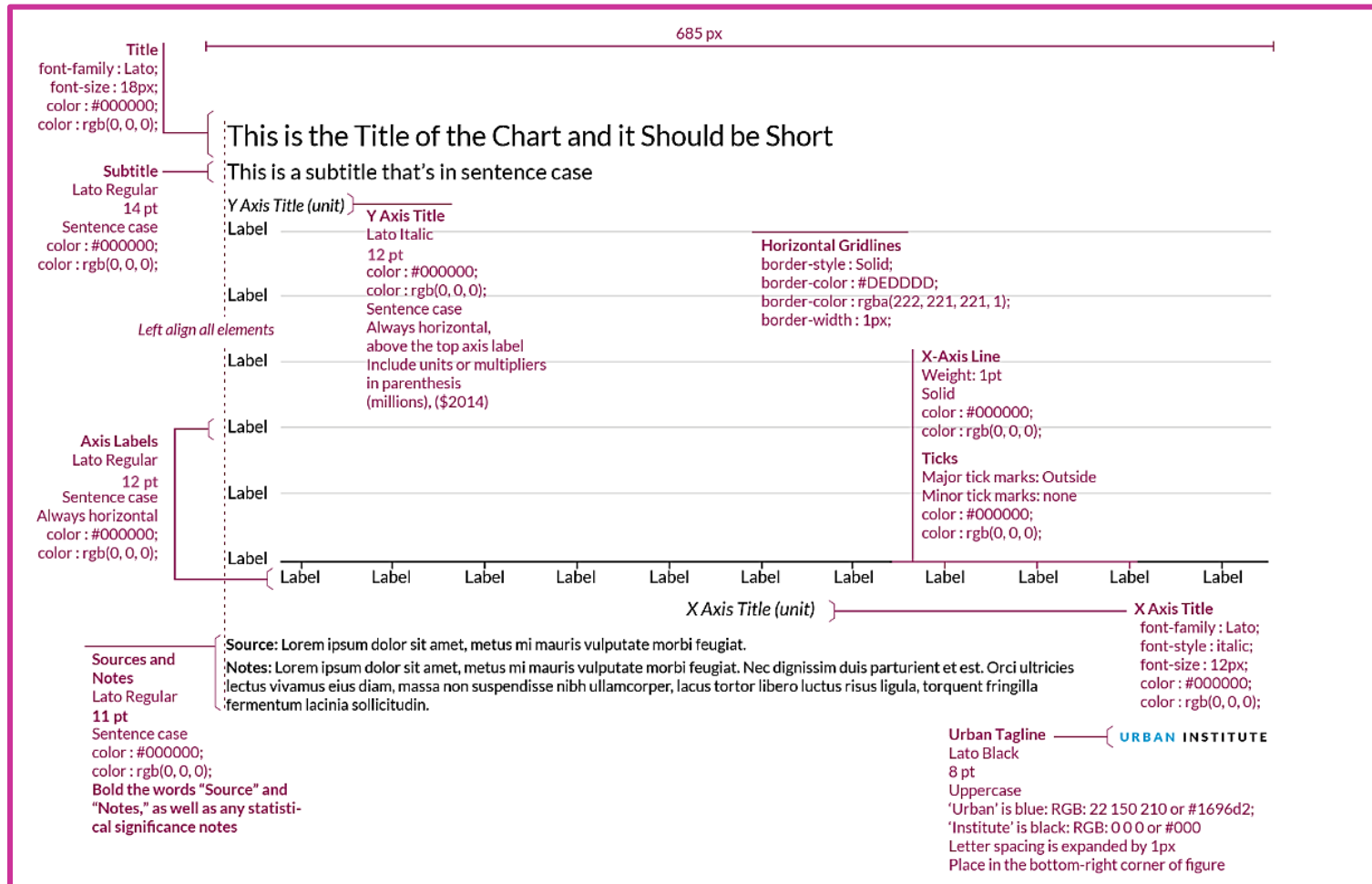
2. Know your data. How you visualize your data will depend in part on the type of data you have. Before deciding on a type of data visualization, you should:

- **Understand what kind of data you have.** Different types of data lend themselves to different types of visualizations. Solid foundational knowledge about the variables your data contains and the data type will help ensure that you choose an appropriate visualization. For instance, it will be difficult, if not impossible, for you to create a map without having location data.
- **Define a clear purpose.** For a static data visualization, figure out what story or message you want to tell based on the data. For an interactive data visualization, think about how you want your users to navigate through the data and what patterns you want them to explore. This will be one of the main driving factors behind the decisions you make about the types of data visualization and elements to include. You will want to make choices that help to reinforce the story or message.

3. Create a data visualization style guide. A style guide provides visual and editorial guidelines to produce a cohesive look and feel across your charts. It will also help to save valuable time by limiting the number of decisions you need to make and ensure that all team members apply data visualization best practices. Exhibit VI.1, below, is an example of a chart style guide.



Exhibit VI.1. Example Data Visualization Style Guide



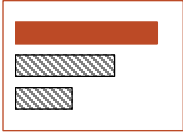

Source: Urban Institute Data Visualization Style Guide. <https://urbaninstitute.github.io/graphics-styleguide/>.

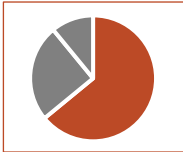






Step 2. Develop compelling content

- 1. Choose the right type of data visualization.** This is a critical choice that may be the most important driver of audience adoption.
 - **Evaluate your choices.** There are three broad types of data visualizations.
 - A **static visualization** refers to a data display that cannot be changed by users. It is predetermined by the author and does not change over time or based on real-time data. Static visualizations typically incorporate no movement. All the information is presented at the same time.
 - An **animated data visualization** reveals data in stages. For example, instead of showing a trend line for all years at the same time, the image moves, revealing additional years and changes in the trend line over time.
 - An **interactive visualization**, such as a data dashboard, allows people to make selections that change what data, trends, and comparisons they see. Users interact typically by clicking, filtering, or doing some type of action that triggers more information to become visible.
 - **Make sure you have adequate resources.** Costs for developing different types of visualization can vary substantially, depending on the software tools, staff, and data that the study team has available. An interactive visualization is likely to be substantially more expensive to develop compared to the first two, in part because user testing is critical for developing that research product.
 - Consider **audience engagement patterns and levels of prior knowledge** when selecting the type of visualization.
 - A static visualization offers greater control of the key takeaway message but may be less engaging. Research suggests that audiences enjoy animated data visualizations more but may leave with a less accurate recollection of the data presented.
 - An interactive visualization allows the audience to explore the data more and develop customized views applicable to their context and use cases.
- 2. Pick the best chart.** Whether your visualization is static or interactive, you will need to decide what type of chart to use. As shown in Exhibit VI.2, there are numerous types of charts from which you can choose.

Exhibit VI.2. Chart Types

| Chart Type | What is it? When to use it? |
|---|---|
| <p>Bar Chart</p>  | <p>Bar charts consist of one or more series of vertical or horizontal bars. Bar charts with vertical bars are also referred to as column charts. There are several types of bar charts, including grouped, stacked, and waterfall charts.</p> <p>Bar charts are one of the most common types of charts. The frequency with which people encounter bar charts gives them an advantage over other types of charts. Bar charts are easier for most people to interpret and understand.</p> <p>Consider a bar chart if you want to:</p> <ul style="list-style-type: none"> Compare information for two or more categories Illustrate comparative rankings Highlight differences between categories <p>Best practices for bar charts:</p> <ul style="list-style-type: none"> Use a zero baseline (i.e., the point at which the x-axis intersects with the y-axis) to ensure that you are providing a visually accurate comparison. Make your bars wide and ensure the gap between the bars is smaller than the bars themselves. Do not include too many categories. Doing so will clutter your chart and make it difficult to interpret. |
| <p>Line Chart</p>  | <p>Line charts connect continuous data points using a line. The term “continuous data” refers to numeric values that are not fixed. The data can change over time and can be broken down into more precise measurements.</p> <p>Consider a line chart if you want to:</p> <ul style="list-style-type: none"> Show changes in data over time Show differences or relationships between multiple categories of information over time <p>Best practices for line charts:</p> <ul style="list-style-type: none"> Use consistent intervals between data points. For instance, if you are showing trends over time, you will want to make sure that you are consistent with the time points plotted. Limit the number of lines you display to 5 to 7 lines. Too many lines can make the chart difficult to interpret and overwhelm your audience. |

| Chart Type | What is it? When to use it? |
|--|---|
| Pie Charts  | <p>Pie charts show information as segments of a circular image. Pie charts are particularly useful for displaying proportions of a specific category.</p> <p>Consider pie charts if you want to:</p> <ul style="list-style-type: none"> Show parts that add up to 100%. Display information so it can be easily eyeballed and does not require precise comparisons. <p>Best practices for pie charts:</p> <ul style="list-style-type: none"> Do not include too many portions. The more wedges there in the pie, the harder it is to decipher. Keep it flat. In other words, do not make it 3D. A 3D pie chart can distort the proportions. This will make it more likely that your audience will misinterpret the information. |
| Maps  | <p>Maps are used to show information by location. This can be done using different colors or markers, such as bubbles.</p> <p>Consider using a map if you want to:</p> <ul style="list-style-type: none"> Focus attention on a specific geographic location <p>Best practices for maps:</p> <ul style="list-style-type: none"> Don't use a map if location is not an important part of the analytic story. |
| Scatter charts  | <p>Scatter charts—also referred to as a scatter plot—are used to show a relationship, trend, or pattern for two numerical variables. Information in this type of chart is depicted as dots. Scatter charts allow viewers to easily understand a relationship or trend by observing how the dots (or data points) cluster together or remain scattered. It is also helpful for detecting possible anomalies in the data. Scatter charts may also include a trend line that is used to highlight the direction of the relationship.</p> <p>Consider using a scatter chart if you want to:</p> <ul style="list-style-type: none"> Show correlation and clustering, especially if you have a lot of data points. Identify possible anomalies or outliers. <p>Best practices for scatter charts:</p> <ul style="list-style-type: none"> Make sure your y-axis starts at zero and the scale is evenly distributed to prevent visual distortion of the data, which could potentially lead to incorrect conclusions. Vary the color of the dots to include additional relevant information to the chart. Include enough data points to show a visible pattern or relationship in the data. But be careful not to include so many data points that they look like a large blob. |

| Chart Type | What is it? When to use it? |
|---|--|
| <p>Bubble charts</p>  | <p>Bubble charts are used to show the relationship between three or more variables, with each data point represented by a bubble. The size of a bubble serves to highlight an additional aspect of the relationship.</p> <p>Consider using a bubble chart if you want to:</p> <ul style="list-style-type: none"> Show the relationship between three or more variables. <p>Best practices for bubble charts:</p> <ul style="list-style-type: none"> Minimize the amount of data depicted. Too much data can make the chart confusing and difficult to understand, as the bubbles may start overlapping or clumping together. Bubbles can only represent data with positive values. It is not possible for bubbles to represent data with negative values. |
| <p>Area charts</p>  | <p>Area charts are a form of a line graph with the area under the line filled in with a color. Area charts are useful for showing trends over time for one or more categories of data.</p> <p>Consider using an area chart if you want to:</p> <ul style="list-style-type: none"> Make comparisons between categories or groups over time. Show the component parts of a whole. <p>Best practices for area charts:</p> <ul style="list-style-type: none"> Minimize the number of categories or groups shown in the chart. Take into consideration the order of the lines if they are stacked or overlap. Use zero as your baseline to prevent distortion of the data. |

Source: Developed by authors.

- I. A Tool Kit Overview
- II. Planning for E2A Products
- III. Written Products
- IV. Presentations
- V. Infographics & One-pagers
- VI. Data Visualization

CHECK OUT THESE RESOURCES TO HELP YOU CHOOSE THE RIGHT CHART

There are many resources available to help you choose the right chart.

- Data Viz Project: <https://datavizproject.com/about/>
- Visual Vocabulary: <https://public.tableau.com/app/profile/andy.kriebel/viz/VisualVocabulary/VisualVocabulary>
- Data Wrapper: <https://app.datawrapper.de/river/>
- Datylon Chart Library: <https://www.datylon.com/product/chart-library>
- From Data to Viz: <https://www.data-to-viz.com/>

3. For interactive visualizations to be effective, you will need to consider the following additional questions:

- **How do you want potential users to move through the interactive chart or dashboard?** Answering this question requires thinking about your potential users' journey through the data and the functionalities needed to provide them with that experience. You will need to think about the sequence of actions users will take and the order (if any) of the charts and information they will see or experience.
- **How much data and how many features should you include?** The answer to this question is particularly important for dashboards. You may be tempted to add as many charts and data streams as possible to your dashboard. But doing so is more likely to detract from than enhance the user experience. You will need to decide what cuts of the data you want users to see and what parameters to use. For instance, if your dashboard involves exploring changes in labor market demand over time, you will need to decide if it is important for users to see the data by year, decade, or some other time interval.
- **What software/platform will you use?** There is a wide selection of software available to support the development of interactive data visualizations. Commonly used software includes Tableau, Power BI, Dundas BI, JupyterLab, D3.js, Shiny, FusionCharts, and Infogram. Each of these has its own strengths and weaknesses that you will need to assess to determine which is the right one for the job.





Step 3: Design for Focus

- 1. Create a coherent design with a visual hierarchy.** The use of a coherent and consistent design will enable users to focus on the information, while the use of a visual hierarchy will make it easier for people to see and find the information that is relevant to them. This approach will also help to ensure that the most crucial data points and messages are prominently displayed. Strategies for creating a visual hierarchy include:
 - **Using colors to highlight and clarify your message.** Contrasting colors can help guide your audience's attention to the key points. To emphasize the most important information, consider using a color that stands out from the rest of the chart. For example, you may want to use muted colors (e.g., gray) throughout the chart and for baseline data, and a vibrant color (e.g., red) for the main insight and/or data point.
 - Intentional use of **font type and size** to guide your audience to the important information.
- 2. Put the key insight in the title.** Instead of a generic title, such as Number of XXX, use a descriptive title that tells the audience the main finding. Your audience may not spend more than 30 seconds looking at your chart. Ask yourself, what is the one thing you want someone to remember? A 6- to 12-word descriptive title will make it easier for your audience to understand and remember the key insight.
- 3. Reduce the noise.** Visual elements that do not contribute to the informative value of the chart create clutter and divert people's attention from the intended focus. The more visual elements there are, the more brain power is required to process the information. This increases the likelihood that people will move on before absorbing the message that the visual is trying to convey.

Below (Exhibit VI.3) is an illustration of how decluttering your chart can help your audience focus on the message you are trying to communicate.

Reminder: See Chapter II. Planning for E2A Products for tips on developing accessible products.



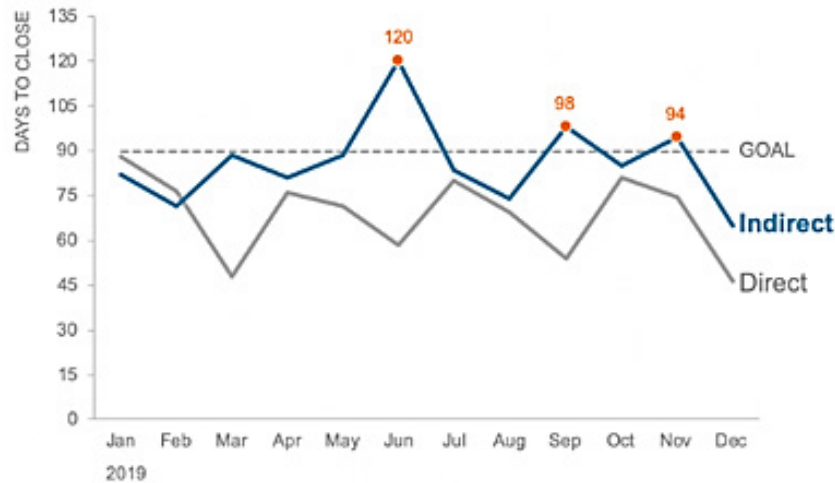
Exhibit VI.3. Decluttering Example

Original graph:



Decluttered & focused:

Time to close deal: **indirect sales missed goal 3 times**



TITLE: Summarizes topic in a bold takeaway in a bright color

CONTENT: Removes repetitive content (year) and distracting data (point estimates for each bar)
 Includes informative axis labels and goal as a dotted line

FORMAT: Removes distracting grid lines, uses bright colors only for key trend and data points

Source: Adapted by authors from: #SWDchallenge: Declutter & focus. <https://www.storytellingwithdata.com/blog/2022/2/1/swdchallenge-declutter-and-focus>.

Methods for reducing visual noise include:

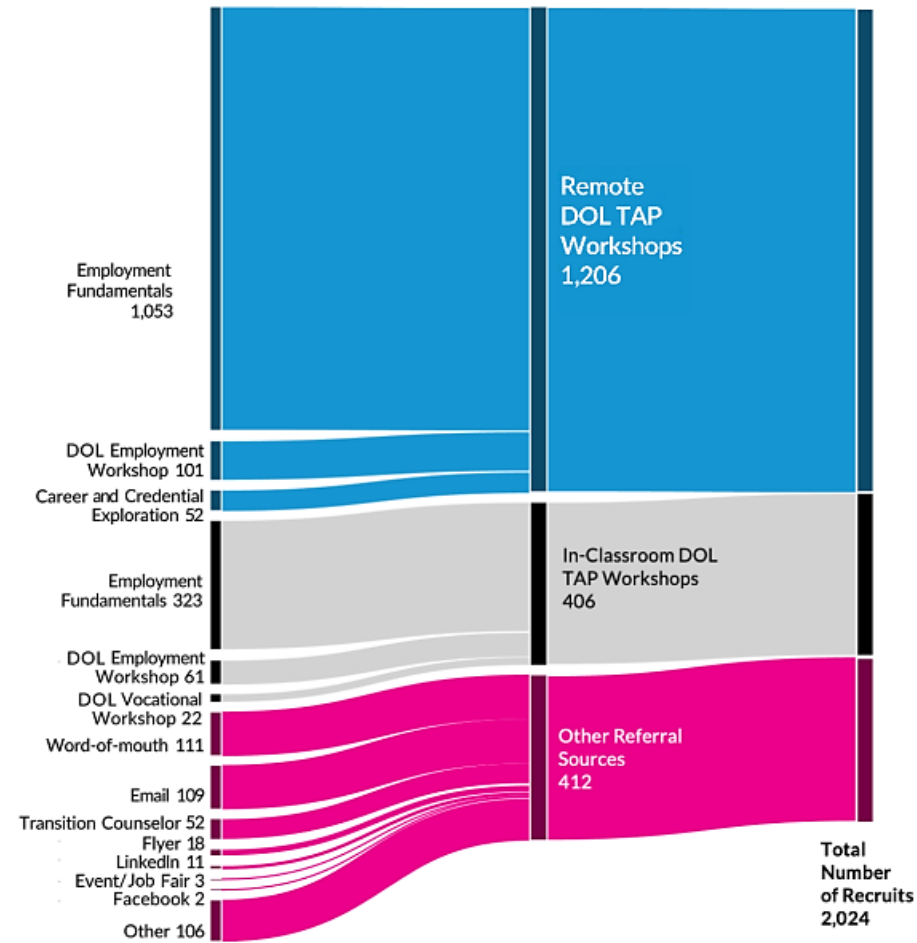
- **Removing the legend and labeling your data directly** – Positioning data labels close to the data helps your audience establish clear connections between them. Legends can cause people to do more work to understand the data visualization by forcing them to move back and forth between the data and the legend.
 - **Deleting grid lines** – if you choose to use grid lines, make them a light gray or another muted color.
 - **Limiting the number of colors** – Too many colors can overwhelm your audience. Instead, choose a limited number of colors and use them intentionally to guide your audience to the main points.
4. **Design for inclusion.** Charts with low contrast can be challenging for people with visual impairments. Use tools that allow you to assess how an image will be visualized by people with visual impairments, such as the color blindness proofing feature in Photoshop and Illustrator.
 5. **Make sure that the proportions are not distorted.** Distorted proportions in a chart can lead your audience to draw incorrect conclusions. For this reason, we recommend avoiding the use of 3D charts.



Additional Data Visualization Examples

Example: Example of Sankey Chart

Figure 3.2: Referral Sources and Number of Recruits for the Apprenticeship Pilot, April 2020 to March 2021

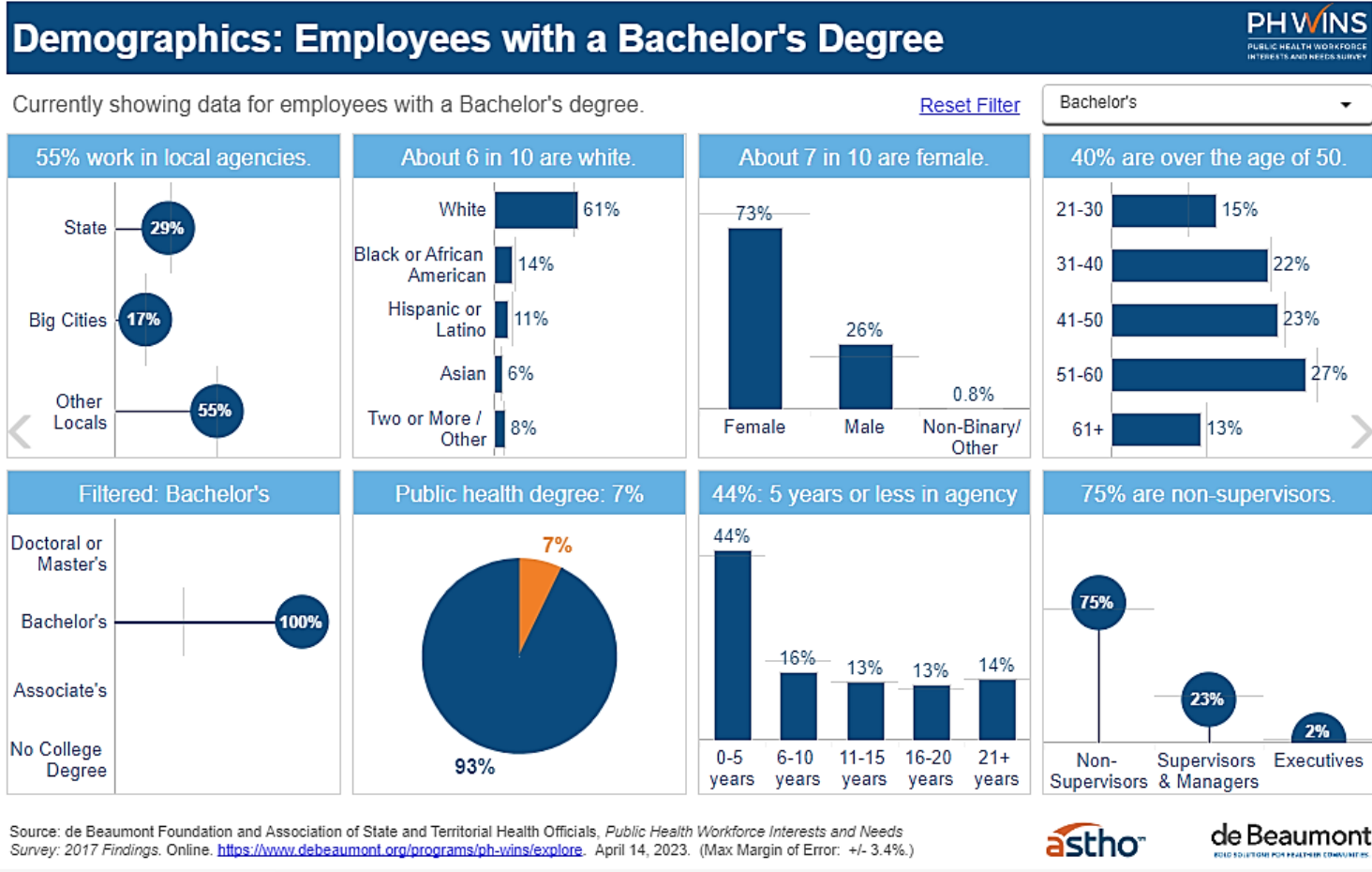


Source: TAP4ME data.

Source: https://www.dol.gov/sites/dolgov/files/OASP/evaluation/pdf/VETSApprenticeshipPilot_ImplementationFinalReport.pdf.

- I. A Tool Kit Overview
- II. Planning for E2A Products
- III. Written Products
- IV. Presentations
- V. Infographics & One-pagers
- VI. Data Visualization

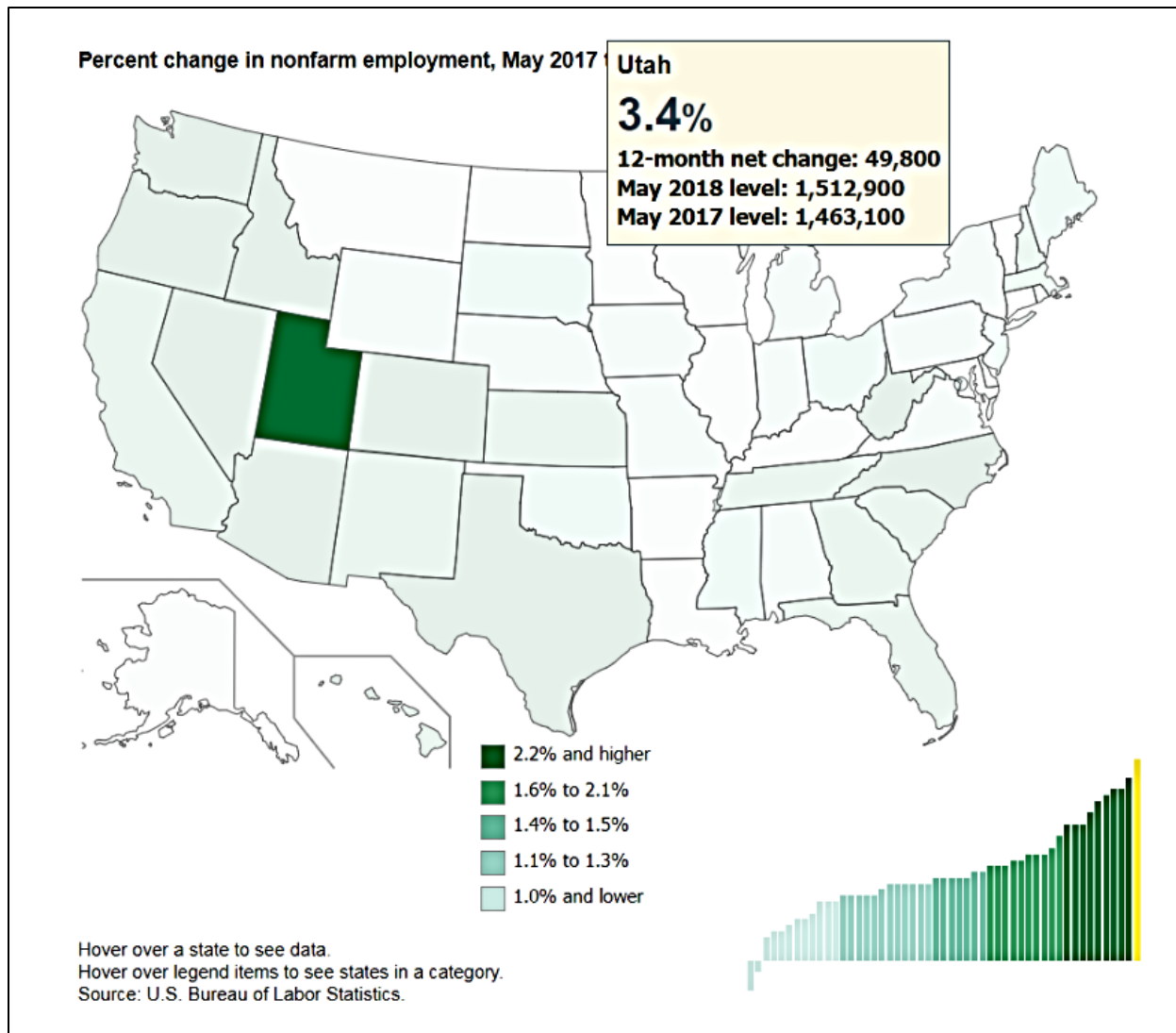
Example: Data Dashboard



Source: <https://debeaumont.org/phwins/explore-the-data/>.

- I. A Tool Kit Overview
- II. Planning for E2A Products
- III. Written Products
- IV. Presentations
- V. Infographics & One-pagers
- VI. Data Visualization

Example: Interactive Map



Source: <https://palewire.github.io/usa-style-guides/bls/data-viz/index.html#maps>.



Data Visualization Checklist

❖ Plan

- Understand your audience's needs.
- Know your data
 - Understand the strengths and limits of the data you have
 - Figure out what story or message you want to tell
- Create a data visualization style guide

❖ Develop compelling content

- Choose the right type of data visualization
 - A static visualization conveys a specific story all at once
 - An animated visualization unfolds a specific story in stages
 - An interactive visualization allow data exploration
- For effective interactive visualizations, spend time on the following questions:
 - How do you want users to move through the interactive chart or dashboard?
 - How much data and how many interactive elements should you include?
 - What software/platform will you use?
- Pick the best chart

❖ Design to Focus Attention

- Create a visual hierarchy using color, font type, and font size
- Put the key insight in the title
- Reduce the noise – remove visual elements that do not have informative value
- Make sure proportions are not distorted – avoid 3D charts

I. A Tool Kit
Overview

II. Planning for
E2A Products

III. Written
Products

IV. Presentations

V. Infographics
& One-pagers

VI. Data
Visualization

Resources and Sources

Data Visualization Design Principles

- Ajani, K., Lee, E., Xiong, C., Knaflic, C. N., Kemper, W., & Franconeri, S. (2021). *Declutter and focus: Empirically Evaluating Design Guidelines for Effective Data Communication*. IEEE Transactions on Visualization and Computer Graphics.
- Evergreen, S. D. (2019). *Effective Data Visualization: The Right Chart for the Right Data*. SAGE publications.
- Knaflic, C. N. (2015). *Storytelling with Data: A Data Visualization Guide for Business Professionals*. John Wiley.
- Franconeri, S. L., Padilla, L. M., Shah, P., Zacks, J. M., & Hullman, J. (2021). The Science of Visual Data Communication: What Works. *Psychological Science in the Public Interest*, 22(3), 110–161.
- Nussbaumer-Knaflic, C. (February 2022). #SWDchallenge: Declutter & focus.
- Robertson, G., Fernandez, R., Fisher, D., Lee, B., & Stasko, J. (2008). Effectiveness of Animation in Trend Visualization. *IEEE Transactions on Visualization and Computer Graphics*. 12(6), 1325-1332.

Equity

- Equity Data Hub. Tableau Foundation. <https://www.tableau.com/foundation/data-equity>
- One Graph or Many: Taking an Equity Lens When Visualizing Data by Race. (2023). Urban Institute. <https://urban-institute.medium.com/one-graph-or-many-taking-an-equity-lens-when-visualizing-data-by-race-bae884dee11c>
- Racial Equity in Data Visualization Checklist. (2021). Urban Institute. <https://www.urban.org/sites/default/files/2021/06/08/do-no-harm-guide-checklist.pdf>
- Schwabish, J., & Feng, A. (2021) *Do no harm guide: Applying equity awareness in data visualization*. Urban Institute. <https://www.urban.org/research/publication/do-no-harm-guide-applying-equity-awareness-data-visualization>



Inclusive & Accessible Design

- The Best Charts for Colorblind Viewers: <https://www.datylon.com/blog/data-visualization-for-colorblind-readers>
- Colour Contrast Checker: <https://colourcontrast.cc/>
- Designing Accessible Data Visualizations: <https://accessibility.huit.harvard.edu/data-viz-charts-graphs>
- Tips for Writing Alternative Text for Images: <https://accessibility.huit.harvard.edu/describe-content-images>

I. A Tool Kit
Overview

II. Planning for
E2A Products

III. Written
Products

IV. Presentations

V. Infographics
& One-pagers

VI. Data
Visualization

About this Project

This toolkit was prepared by the American Institutes for Research for the U.S. Department of Labor, Chief Evaluation Office under the CEO Administrative Data Research and Analysis contract (DOL # 1605DC-18-A-0018), which is managed by Insight Policy Research. The views expressed are those of the authors and should not be attributed to DOL, nor does mention of trade names, commercial products, or organizations imply endorsement of same by the U.S. Government.

Suggested Citation

Amin, S., Corea, C., Souvanna, P., and Mills De La Rosa, S. (2023). *E2A Tool Kit: Tips for Developing Great Evidence to Action (E2A) Products*. Toolkit submitted to the U.S. Department of Labor Chief Evaluation Office. Arlington, VA: American Institutes for Research.

Acknowledgments

The authors are grateful to Megan Lizik and Janet Javar at the Department of Labor for their championship of this effort, and their leadership and guidance on this project.

We would like to thank many people for their excellent inputs, review, and/or feedback, including staff from the Chief Evaluation Office and Chief Data Office at the U.S. Department of Labor; the Administrative Data Research and Analysis team at Insight Policy Research; the Capacity Assessment project team at AIR; the Reemployment Services and Eligibility Assessments Evaluation Technical Assistance Teams at Abt Associates and the Urban Institute; the Reentry Projects and America’s Promise project teams at Mathematica; and the Evaluation Support team at AVAR Consulting.

Finally, we are grateful for the scores of experts from across disciplines who share their ideas with the world on how to communicate effectively and influence action. We are learning from you, applying your recommendations, and learning and adapting as we go. We hope this collation of the many ideas you’ve shared and additions to them do justice to your wisdom. All errors are our own.

Westat[®] insight



Advancing Evidence.
Improving Lives.



Chief Evaluation Office
U.S. DEPARTMENT OF LABOR 