Psychological Science in the Public Interest
2021, Vol. 22(3) 97–109

DOI: 10.1177/15291006211057899
www.psychologicalscience.org/PSPI
SAGE

The Practice of Visual Data Communication: What Works

Jonathan Schwabish
Urban Institute, Washington, DC

The practice of data visualization is both a science and an art. There is science behind how humans' eyes and brains process visual content, and statistical methods behind collecting, processing, analyzing, and preparing data to generate graphs, charts, and diagrams. But the *art* of data visualization is how we bring people into the visual, how we engage them, and how we make them care about the content we are communicating to them.

The target article of this commentary, the thorough work by Franconeri et al. (2021), sets the stage to understand the academic underpinnings of data visualization—how people's eyes and brains facilitate the understanding of visual content, how to design perceptually efficient and understandable visualizations, and how people use different platforms and technologies to interact with data and visual content.

The *practice* of data visualization goes further than many of these concepts to consider how data are plotted, how to use colors and fonts, and how to facilitate engagement and understanding. The standard graphs that many of us have come to know and create, such as line charts, bar charts, and pie charts, are familiar to most readers and easy to read. But many other graph types can be used to communicate ideas and arguments. In the "How to Design an Understandable Visualization" section of their article, Franconeri et al. briefly discuss four alternative graph types (or what I call *nonstandard* graph types): connected scatterplot, parallel-coordinates plot, tree map, and node-link diagram. There, the authors focus on how people in specific fields use specific graphs—for example, engineers and economists use connected scatterplots—not the potential for these formats to engage audiences on a broader level. In some cases, such nonstandard graph types can be inherently better at communicating data and in other cases are simply more engaging, which can be a goal in and of itself. Whether you are a researcher, analyst, marketer, or journalist, you know that the amount of content people see every day makes grabbing and maintaining attention difficult; thus, *engagement* can be a crucially important part of the data communicator's toolkit.

Here, I present several alternatives to the standard ways of visualizing and communicating a relatively simple data set from the *National Center for Education Statistics* (NCES). From my perspective, these alternative graphs are not so far outside the experience of most readers that they cannot be used more frequently—in the language of Franconeri et al., the "schema" in these graphs are well known and consist of dots, lines, and icons. The goal of this commentary is not to argue that the presented graphs are somehow the "best" that can be created with these data. Instead, my goal is to demonstrate the array of visual options we have to communicate data and how those options enable us to highlight different patterns or values, and to draw out our own stories for readers and help them reach conclusions.

Five Principles for Better Data Visualizations

To create better, more effective visualizations, I find it useful to follow five basic principles, which I explore in more detail in my book, *Better Data Visualizations* (Schwabish 2021a).

First, *show the data*. People are reading the graph, chart, or diagram to learn something about your data or your argument. The data are the most important part of your visuals and should be shown in the clearest way possible. This does not mean you need to show *all* your data *all* the time, but instead it means being strategic and purposeful about what you want people to see in your visualization. For example, you might have a line chart that shows changes in the high school

Corresponding Author:
Jonathan Schwabish, Urban Institute, Washington, DC
Email: jschwabish@urban.org

graduation rate in each of the 50 states in the United States—instead of showing all 50 lines with different, equally saturated colors, consider using a gray color for all but a few states you want to highlight (also see the latter half of the "How to Design a Perceptually Efficient Visualization" section in Franconeri et al.).

Second, *reduce the clutter*. Chart clutter—heavy gridlines, tick marks, data labels, 3D effects, and more—makes your visualizations less effective and harder to read. The fewer nondata elements you can have on your graph, the easier it will be for your reader to focus on the most important parts of your visuals. This does not necessarily mean stripping away *all* the color, shapes, and icons, or what Edward Tufte and his acolytes derisively refer to as "chart junk." Franconeri et al. make clear that these calls have not been supported by the existing body of research: "Despite strong calls to declutter visualizations (e.g., Tufte, 1983), there is only mixed evidence that this practice improves aesthetic ratings and little evidence that the prescription affects objective performance" (p. 131).

Third, *integrate the graphics and text*. Too many data communicators—especially researchers writing reports and briefs—treat the text and the graphs as two separate elements. The text is used to fully explain the concepts and make the argument, and the graphs are often there to "break up" the text or to show some basic pattern. But the visuals are an integral part of the argument, not separate, so integrate them all: Directly label the lines, bars, and circles on your charts instead of using separated legends; use concise, active titles to tell the reader what they should learn from the graph instead of simply describing what is in the graph; and add important annotations and labels to highlight important aspects of the graph.

Fourth, *use a small-multiples approach*. This is a technique you can use to break up dense, complex graphs into smaller, repeated multiples. Especially in this modern digital-first society, you should not feel as if you need to pack all your information in a single graph or presentation slide. Instead, break it up into smaller multiple graphs that have similar axes, colors, and layouts so the reader can more easily scan from one graph to the next.

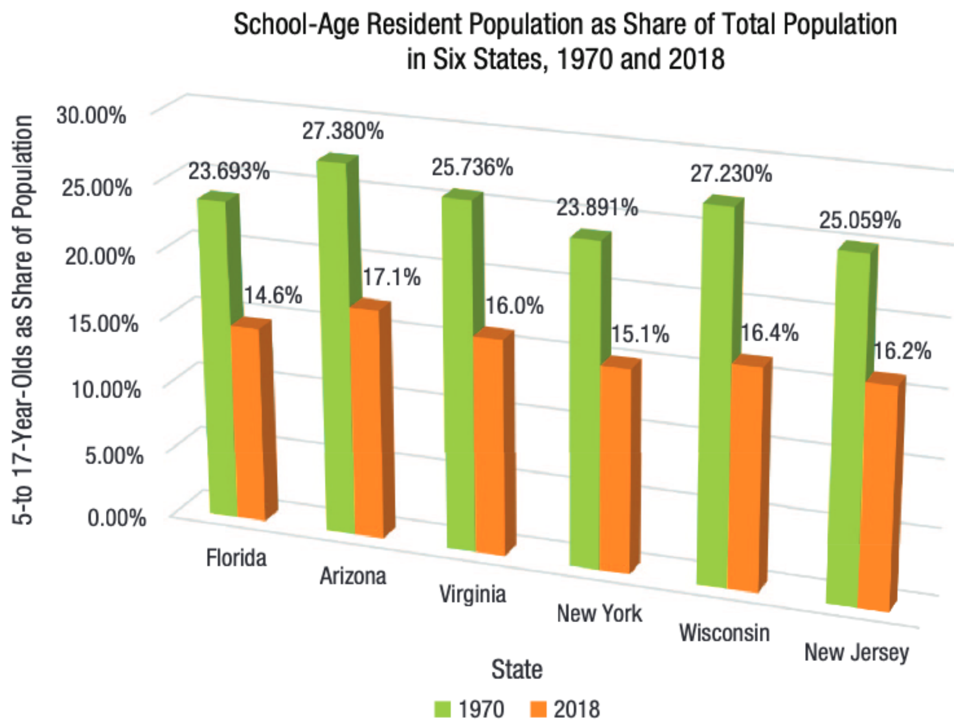
Fifth, *start everything with gray*. Another strategy I use in my visualization-creation process is to try to make everything in my graphs gray. This makes every graph element—from axis to tick mark to bar—equal in terms of how we read and perceive it. With that equal treatment, it forces me to decide where I want to focus the readers' attention, so I might change the thickness of a line, change the color of an outlier point, or make certain bars a different color to make those values easier for the reader to find and read.

These guidelines are secondary to the most important part of visualizing data—*know your audience*. Who is your audience? What do they know about the data, the content, and the graph type you are using to present the data? What level of statistical or data expertise do they bring to your work? A graph published in an academic, peer-reviewed journal article will likely need less explanation of basic data and statistical principles than might a graph published in a conference poster, in an issue brief, on social media, or in a blog post. Always consider who your audience is, what they expect, and how you can best help them reach their (or your) goals through purposeful use of the data visualization principles laid out here and in the Franconeri et al. article.

All five of these strategies are employed throughout the examples below, but first look at a basic bar chart that desperately needs to be redesigned to be more effective (Fig. 1, top). You have probably seen your fair share of graphs that look like this (perhaps you have even created some yourself!). This graph shows the school-age population (5- to 17-year-olds) as a share of the total population in six states in 1970 and 2018. We have the gratuitous 3D effect, which only serves to distort the data (notice how the top of the 15% bar for New York in 2018 does not hit the 15% gridline even though the value is 15.1%), the rotated *x*-axis labels that are difficult to read, the legend at the bottom disconnected from the content, and numerous and inconsistent numbers of digits throughout. The title of the chart simply describes the graph and really does not say much about what people are supposed to learn from the graph.

This graph needs to be redesigned to make it more effectively communicate the content (Fig 1, bottom). First, the 3D effect has been removed and the data are now sorted from largest to smallest; both techniques make the graph easier to read. The *x*-axis labels are horizontally oriented, and the legend and title are integrated at the top left of the graph. The data labels are now consistent, with one digit each, and I have removed the gridlines and vertical axis labels, which are unnecessary with the data values placed above the bars.

Also notice the change in color from the previous graph. Color selection is a big challenge in the data visualization field, especially because many of us (me included) are not graphic designers and prefer to leave those decisions to experts. One of the more important things to consider when it comes to color is that about 4% of the global population (Olson & Brewer, 1997) has some form of color-vision deficiency (often called "color blindness"), the most common of which involves difficulties discerning between similar shades of green and red. Notice what happens when basic rainbow colors are adjusted to mimic color-vision deficiency in



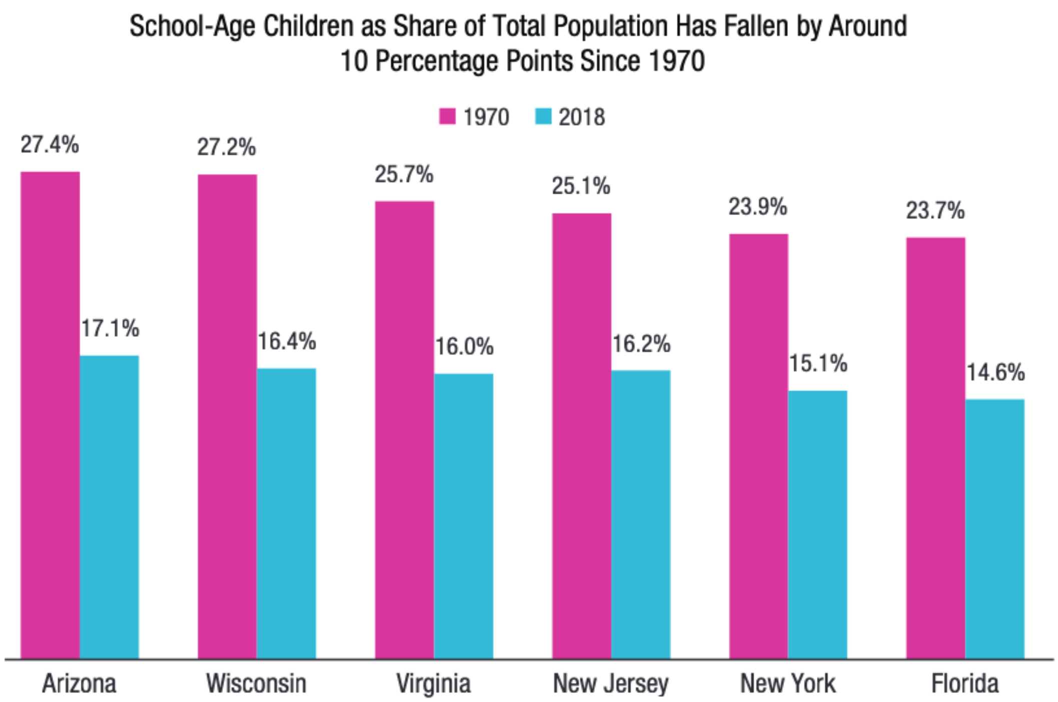


Fig. 1. Example of redesigning a bar chart to communicate the data more clearly and effectively by removing unnecessary elements, making numbers more consistent, and employing a more consistent color palette. School-age children are defined as 5- to 17-year-olds. Data source: National Center for Education Statistics (2018).

Figure 2—it becomes almost impossible to distinguish some of them.

There are numerous resources you can use to help choose your color palettes—such as Color Oracle (https://colororacle.org/), Coblis Color Blindness Simulator (https://www.color-blindness.com/coblis-color-blindness-simulator/), and the Color Contrast Checker at WebAIM (https://webaim.org/resources/contrastchecker/)—as well

Fig. 2. Example of what someone with the most common form of color-vision deficiency (deuteranopia) might see for these sample colors.

as graphic-design websites and other data visualization style guides (Cesal, 2020; Schwabish, 2016, 2021c, 2021e).

It is also worth noting that although many people working in the data visualization field spend an inordinate amount of time focusing on making accessible color palettes (Franconeri et al. devote a few paragraphs to color-vision impairments; p. 120), less focus is paid to making content accessible to people with other forms of vision impairments, not to mention those with other physical or intellectual impairments, or even those who have difficulty accessing digital content because they live without access to high-speed Internet (Dornauer & Bryce, 2020). Fortunately, a growing body of academic research and practical strategies is being developed to improve how data visualizations can be made accessible to everyone. In the meantime, the GitHub repository from Le Gassik et al. (n.d.) is a good starting point for a series of resources and references about data visualization accessibility.

Overall design features are always important to help communicate visual information, but even more so in nonstandard graphs with which users may be unfamiliar. For example, Franconeri et al. (pp. 42–43) refer to two visualizations that used a counterintuitive inverted *y*-axis (starting from zero at the top and moving to larger numbers down the page) to depict deaths over time during the war in Iraq (Scarr, 2011) and gun deaths in Florida (Christine Chan in Engel, 2014). But what Franconeri et al. omit from their discussion is that the former visualization was a celebrated graphic, winning awards at two major design and news conferences (Malofiej and the Society of News Design; see Scarr, n.d.), whereas the latter was roundly ridiculed on social media and elsewhere for misleading readers (e.g., Lallanilla, 2014).

You can make your own judgment of the differences in two graphs that mimic the design features in the Scarr graphic (Fig. 3, left) and the Chan graphic (Fig. 3, right),

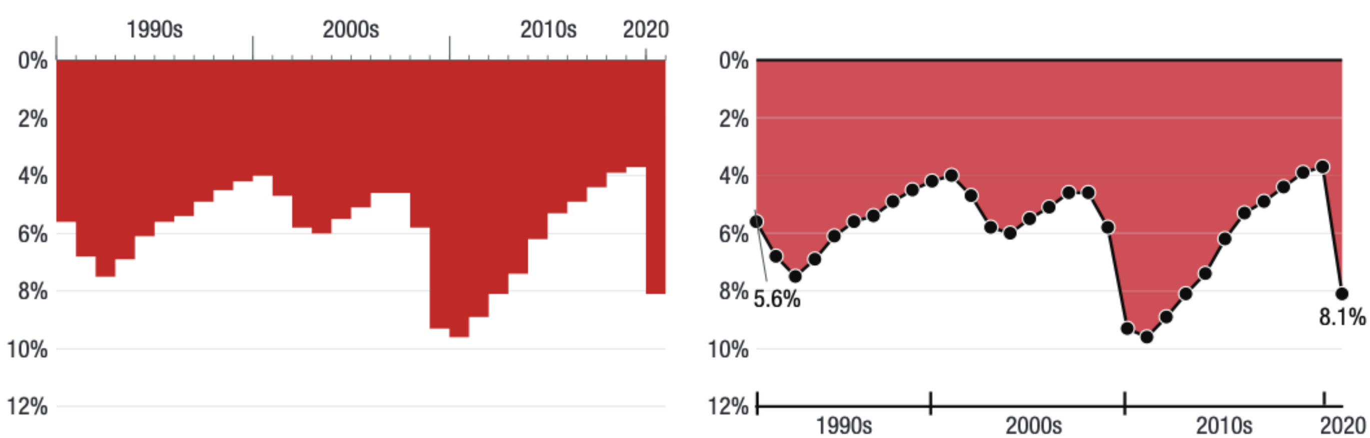


Fig. 3. Demonstration of how design decisions can affect perception of the data. Inspired by Scarr (2011; left) and a graph by Christine Chan in Engel (2014; right).

here using the U.S. national unemployment rate from 1990 to 2020 (U.S. Bureau of Labor Statistics, n.d.). Both graphs use an inverted vertical axis. The graph on the left is a vertical bar chart with horizontal axis labels at the top of the graph aligned with zero. On the right, the same data are shown using an area chart with a black line and circles along the bottom edge, and the horizontal axis labels set below. The difference? The version on the right makes it appear as if the *black line* is encoding the data so that, for example, it looks like the unemployment *rose* in the 2010s and the spiked *downward* in 2020. By comparison, the bar chart on the left—while not necessarily as intuitive as a standard bar chart—focuses attention on the red bars as encoding the data and leaving the white area as empty space. In the framework of the target article, this difference corresponds to focusing the viewer on either the length of the bars (an appropriate encoding used in the graph on the left) or the position of the edge of the line (a misleading approach used in the graph on the right).

Numerous Ways to Visualize Geographic Data

To demonstrate the array of graphic types available to data communicators, I use a simple data set from the NCES (2018), the primary federal agency that collects and analyzes data related to education in the United States and other nations. The data include the average starting time of public schools in each state across the United States. For each U.S. state, the data includes the overall average start time and the distribution of schools within 29-min bands (i.e., 7:30 a.m.–7:59 a.m., 8:00 a.m.–8:29 a.m., and so forth). Here, I focus on just the average starting time and create everything in Microsoft Excel—not because I think Excel is the only or best software tool but to demonstrate that Excel, like many software tools, can be extended in ways to create many graphs and charts not available in the basic drop-down menus.

The data are presented in a dense table on the NCES website. There is nothing inherently wrong with presenting the data in table format, and for those of us who want the data to explore and conduct analyses, providing it in Excel or CSV files rather than buried in PDF files is obviously preferable. But to help more website visitors use and understand the data, providing a graph or visual is beneficial to improving understanding and enhancing engagement.

Maps, Maps, Maps

Most people's initial inclination when being presented with geographic data is to create a map. Although data-driven maps can be useful to show geographic patterns, there is a trade-off between the familiarity of the map and how the geographic areas can distort the importance of the data. Take the United States, for example—the land area of Wyoming is 97,818 square miles, whereas Massachusetts is 10,565 square miles, about one tenth the size. Wyoming has three electoral votes in U.S. presidential elections compared with 11 for Massachusetts. For Presidential elections, which state is more important? Thus, we should ask ourselves whether a standard election map in which states are colored red and blue is the most useful representation of the data.

The message here, then, is to carefully consider whether a map helps support your point or whether another graph type might be better suited to helping your reader understand the data. Are you inherently telling a geographic story? What is your focus and what do you hope your reader will glean from your visual representation of the data?

Take this "choropleth" map of the statewide average starting time data in public schools (Fig. 4, top). (I created this using the relatively new mapping tool in Excel, although I typically prefer to use other tools for maps such as the R programming language, which is more flexible and provides more options.) I used a continuous color palette here, so the green colors fade from the lightest green for the states that start the earliest to the darkest green for those that start the latest. Overall, you can see a basic geographic pattern: States in the southern part of the United States tend to start a little earlier in the day relative to other parts of the country, especially states in the northern and western parts of the country.

This standard choropleth map is not the only kind of map we can create, however. On the bottom of Figure 4 you can see a "tile grid map" in which each state is presented as an equally sized square (also created in Excel). Again, there is a trade-off: The tile grid map is not as familiar as the standard choropleth map shown above, but it might be more engaging for your reader. This approach also enables us to add more data to the chart, such as a line for each state, which can work here because each state is the same size and shape (see, for example, the chart from Kerrie Vila and Chris Canipe in Allen, 2018). Also notice that I converted the standard map legend into a bar chart to show the number of observations in each category. Unlike the continuous color palette used above, this "histogram-as-legend" technique enables your reader to get an immediate sense of what we might think is a basic question: How many states start after 8:24 a.m. or before 8:00 a.m.?

There are many other types of maps to present geographic data—cartograms, dot density maps, proportional symbol maps, hex grid maps—each with its own

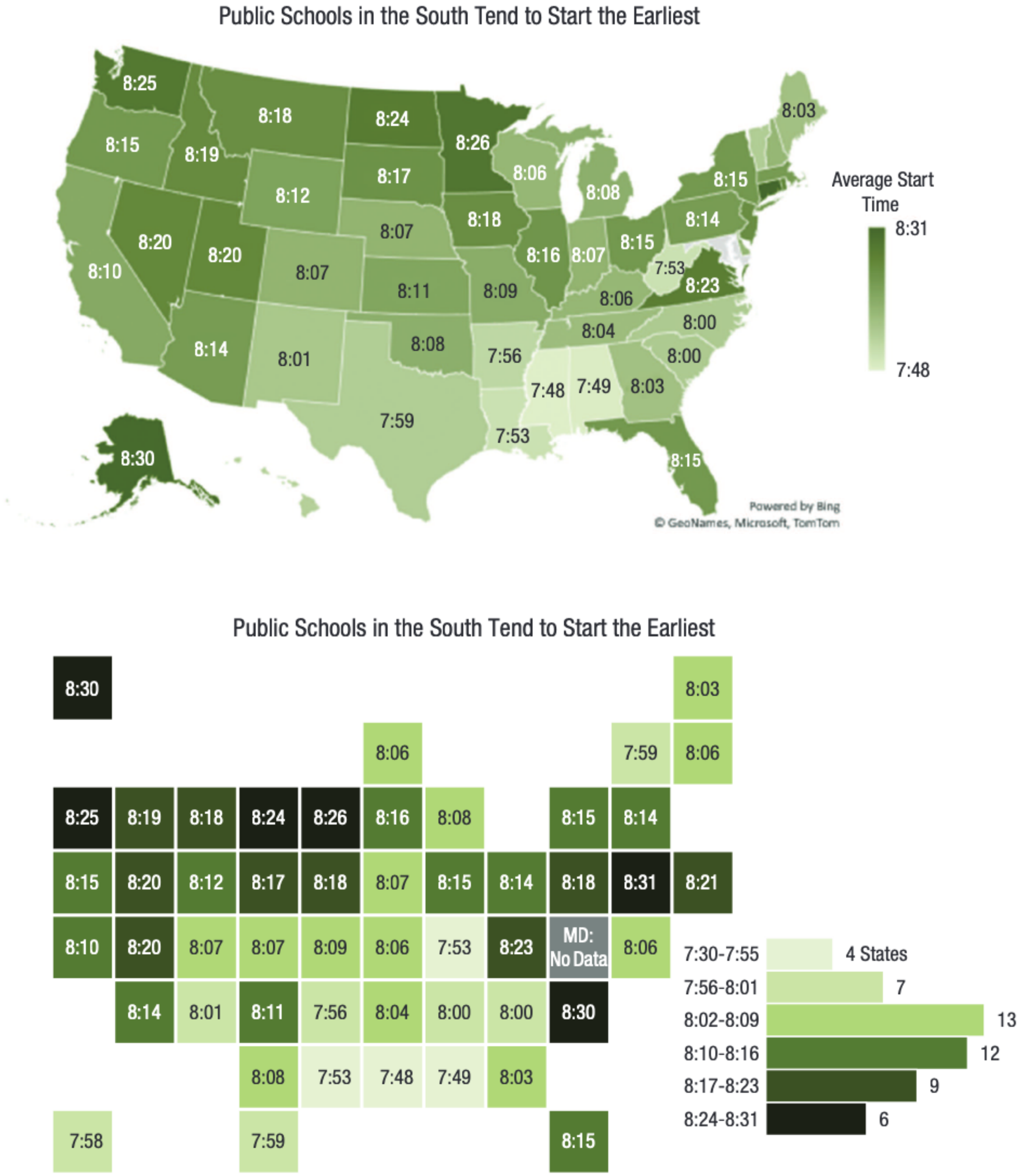


Fig. 4. Two sample maps—a choropleth map (top) and tile grid map (bottom)—to show the average starting time of public schools around the country. Data source: National Center for Education Statistics (2018).

advantages and disadvantages. Again, there is a tradeoff between familiarity and accurately representing the data—which side of the scale you favor will help determine which map will be most useful to you and your reader. Kenneth Field's books *Cartography* and *Thematic Mapping* are perhaps the best books to learn more about data-driven maps (Field, 2018, 2021).

The Ol' Standby: The Bar Chart

The bar chart is another standard way to visualize these kinds of data—a single data point for each state that can be encoded with a bar, which makes it easy to compare the states. As Franconeri et al. note, humans are very good at discerning the quantities from bar

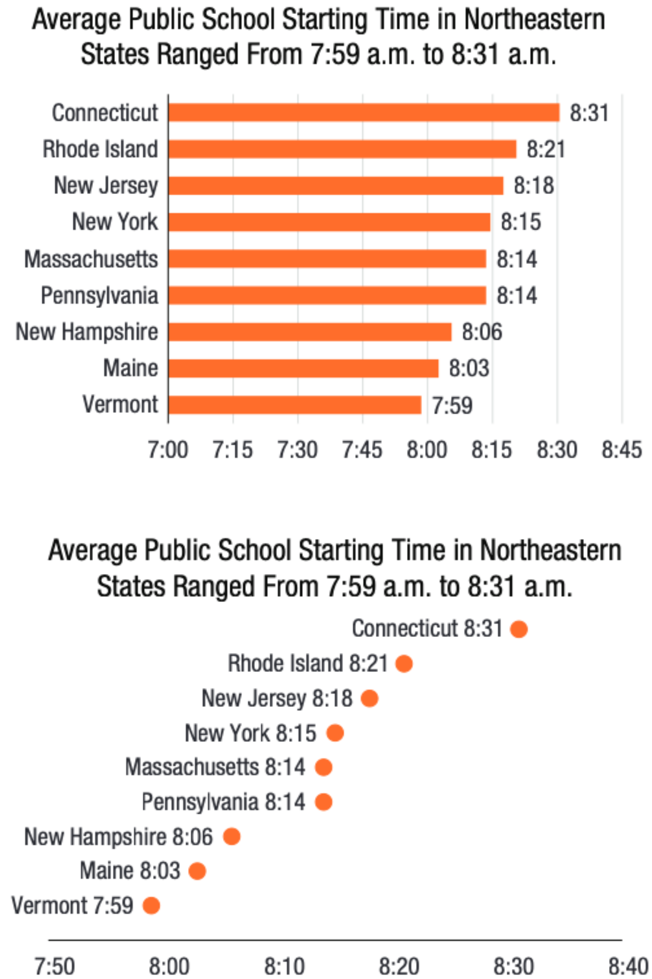


Fig. 5. Both a bar chart and a dot plot can be used to show data, and each has various advantages and disadvantages. Data source: National Center for Education Statistics (2018).

charts when the bars are aligned along the same horizontal axis (see also Cleveland & McGill, 1984, 1985; Heer et al., 2010).

Bar charts can be visually "heavy," however; they can have a lot of "ink" on the page, which, designers argue, can make them difficult to read and digest (Hagen & Golombisky, 2013). In Figure 5, I show the average starting time for just the nine states in the northeastern part of the United States. This is not a huge number of data points, but including all 50 states would make the graph difficult to read because it would be too tall to fit on the page of this journal without making the bars very thin and the text very small. A dot plot (Fig. 5, bottom) is one alternative to presenting data in this way. The dot plot (sometimes called a dumbbell chart, barbell chart, or gap chart) is one of my favorite alternatives to the bar chart primarily because it lightens the space, making additional areas available for labels and annotation.

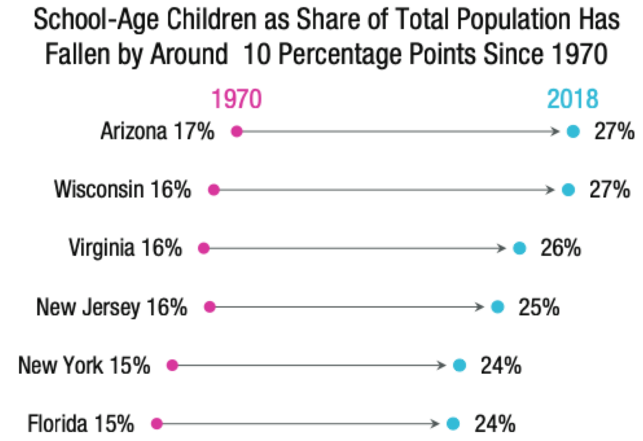


Fig. 6. An example of a dot plot used to show change over time for a sample of states. School-age children are defined as 5- to 17-year-olds. Data source: National Center for Education Statistics (2018).

Dot plots can be especially valuable as an alternative to the paired bar chart in which we ask the reader to make several comparisons simultaneously: both level and change *within* and *across* groups (i.e., states; see Fig. 7 in Franconeri et al., p. 124). In dot plots with two or more dots, a line or arrow can be used to connect them, which in many cases makes it easier to make both kinds of comparisons. The dot plot in Figure 6 uses the same data as in the first set of bar charts that shows the share of the school-age population in 1970 and 2018. Note the use of the year labels directly above the first pair of dots and the color used to further integrate the labels and the data. The arrows help reinforce the concept of time; I am more likely to use a straight line if I am comparing distinct groups.

Spreading Out: Visualizing Distributions

The latter part of the Franconeri et al. article discusses the challenges communicating uncertainty and risk (pp. 139–149). Personally, I would add challenges around communicating *distributions* to this list. Many people simply do not understand statistical concepts such as percentiles, variances, and standard deviations (see, e.g., Wilson, 2018), so the challenge in communicating distributions is often clearly explaining these concepts before conveying the argument or story.

For the school data, we can create a standard histogram that shows the distribution of starting times across the country (Fig. 7). Again, the active title helps direct the reader's attention to the first bar on the left of the graph. We could also change the color of that bar or add additional text annotation to further reinforce the statement in the title.

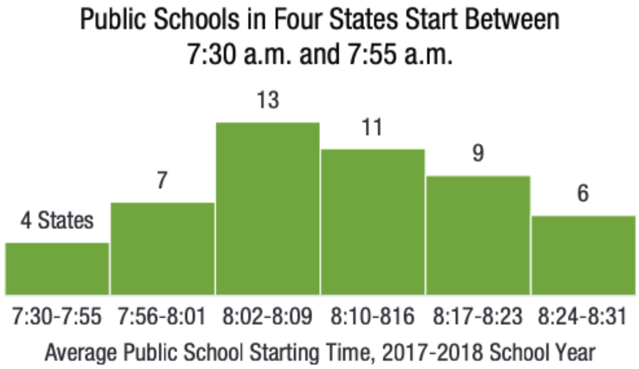


Fig. 7. At its core, a histogram is a vertical bar chart and is used to show the distribution of the data, in this case across six categories or "bins." Data source: National Center for Education Statistics (2018).

One of the challenges inherent in the histogram—and, for that matter, many graph types that use abstract aggregating shapes such as bars and lines—is that those shapes do not necessarily help the reader *connect* with the data. Showing the individual data points can be a useful way to improve the ability of readers to humanize the data by reinforcing that they are looking at people and not just numbers or statistics (see Boy et al., 2017; Groeger, 2014; Lupi, 2017; Morais et al., 2020; Schwabish & Feng, 2021). In this specific example, one approach is to show the actual state names in each bin of the histogram as shown in Figure 8 (top); we can now clearly see where each state falls in the distribution and the use of color further reinforces how southern U.S. states tend to start earlier in the morning relative to the rest of the country.

This technique can be taken a step further by using icons for each state instead of the state name (Fig. 8, bottom). The icon version is fun, but it is probably too difficult for readers to quickly identify specific states—Colorado and Wyoming, for example, have similar shapes. Although icons can be useful in visualizing data and helping readers connect with the data, incorrectly scaling them (see the final section in the Franconeri et al. article) can be problematic. More importantly, icons are often *representations of people*, so as content creators, we need to be careful and mindful of how we are representing different characteristics of people along lines of race, ethnicity, gender, sexual identity, and other characteristics and their intersections (D'Ignazio & Klein, 2020; Schwabish & Feng, 2021).

Correlation and Causation: Visualizing Relationships

It is likely the case that far too much research in the social sciences (and probably elsewhere) assigns a causal interpretation to relationships between variables that are simply correlations. The statistical concepts and tests that can be used to distinguish between correlation and causation are beyond the scope of this article, but it is always worth carefully considering which story you are trying to tell.

The standard way to visualize correlations and associations is the scatterplot—one axis corresponds to one variable and the other axis to another variable. In the scatterplots in Figure 9, I pulled the two data series that I have been using thus far together to show the share of the student age population along the vertical axis and average public school starting time along the horizontal axis.

I added various pieces of annotation here to simply demonstrate how this might be done:

- The nationwide average is denoted with two lines and labels;
- Small annotations help the reader understand the areas that are above and below the U.S. average (along the vertical axis) and schools that start earlier or later (along the horizontal axis); and
- Labels for specific states that are obvious outliers from the rest of the country.

I might not include *all* these labels in a final version and likely not even include the axis-explainer labels if the graph was being published in an academic journal article because I am confident those readers are familiar with how to read a scatterplot. Remember, the audience is of crucial importance when it comes to visualizing data—what do you expect your reader, user, or audience member to know about the content and the graph type you are using?

There are many ways to incorporate many of the techniques discussed thus far into the scatterplot. Icons is a possibility, as are state abbreviations, as shown in the bottom of Figure 9. What is most useful for your audience to both inform and engage is part of the challenge of visualizing data and communicating content.

Sit Down and Take a Load Off: Tables for Visualizing Data

It is perhaps the visualization type that is used the most and receives the least amount of attention: Tables. Of the nearly 270 citations in Franconeri et al., only *one* (Tait et al., 2010) includes the word "table" in the article title (and that article does not deal with *effective* data tables—and the tables in the article itself could use some improvement!). Researchers, especially, are inclined to put their data in tables, likely not only because showing detailed numbers is transparent, but also because tables are easy to create and require very little thought—drop the data

Fig. 8. An alternative to the standard histogram that shows the individual state names or icons instead of the more abstract bar shapes. Both were created in Excel, and the version on the bottom uses the open source StateFace font from ProPublica (https://propublica.github.io/stateface/) to turn letters into state icons. Data source: National Center for Education Statistics (2018).

in, maybe add some boldface to the column headers, add some asterisks to the statistically significant results, and it is ready to go. Aside from the citations in Franconeri et al., there is a body of research on the effectiveness of tables and how to make them better, clearer, and, in some cases, more visual (see, for example, Few, 2004;

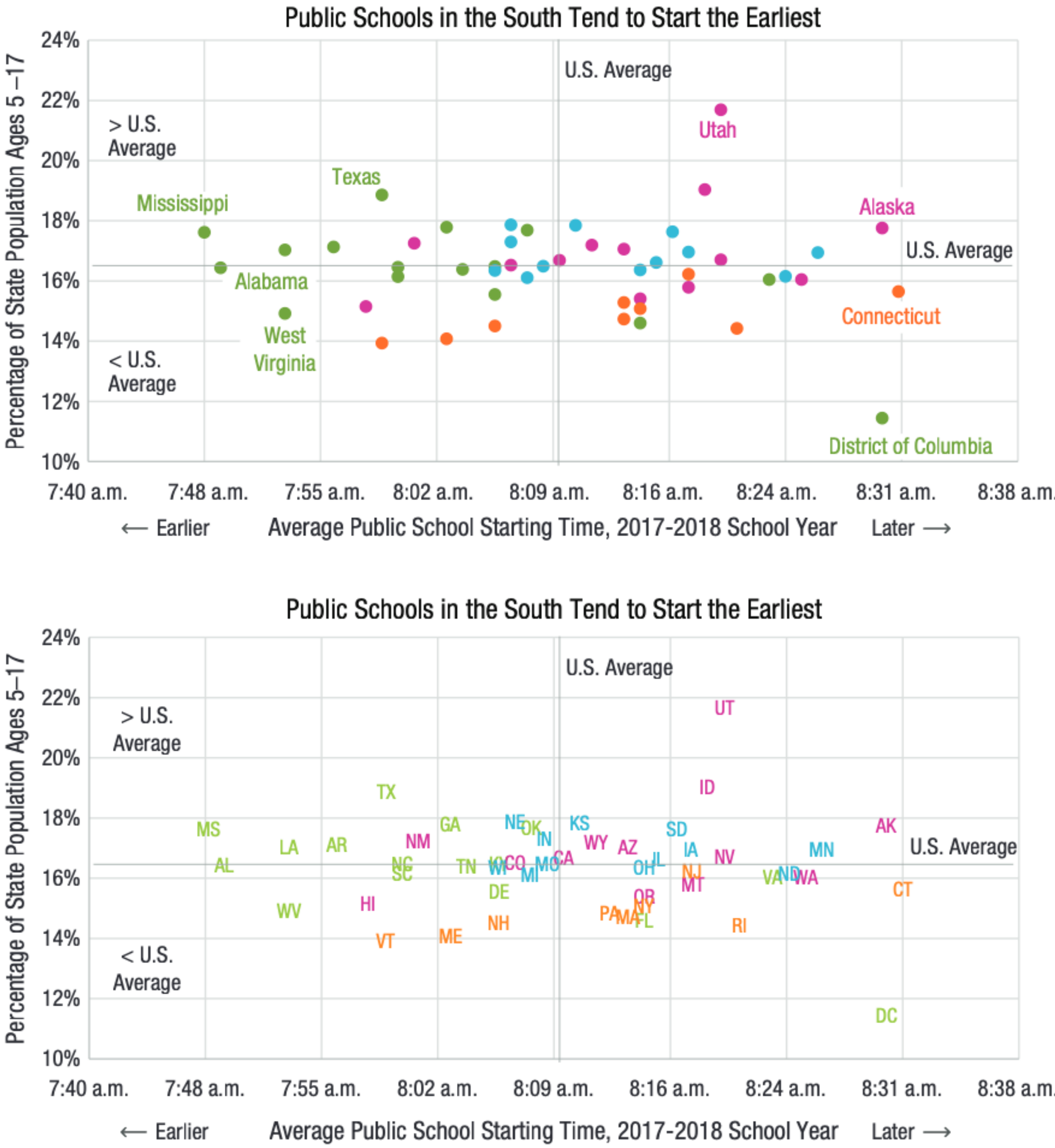


Fig. 9. Two examples of a scatterplot to show the relationship between average public school starting time and the percentage of a state population ages 5 to 17.

Gelman, 2011; Hendel, 2021; Kastellec & Leoni, 2007; Schwabish, 2020).

Tables are useful when you want to show the exact values of your data or estimates. They are probably not the best solution if you want to show a lot of data and are usually not the best solution in a presentation in front of an audience. (It is undeniably true that the worst thing you can say to your audience when putting a dense table of regression coefficients up on the screen is, "I know you can't see this, but . . ." and continuing as if you did not just insult everyone in the room by having just stated you could care less about whether they can actually see the content you are presenting. But I digress. . . .) With a well-designed table, you can guide your reader to specific numbers, patterns, and outliers.

Here is an example from data pulled together from the two datasets used thus far. Perhaps the first table in Figure 10 is an extreme form of bad table design that you might have come across, but it is probably not *that* far off. There are several problems with this table:

- The column headers are not differentiated from the body of the table;

State	Total population 1970	Total population 2018	Students 1970	Students 2018	Percent 1970	Percent 2018	Average start time	% Before 8 a.m.
Arizona	1,775.00	7048.876	486.0	1202.589	27.38%	17%	8:14	29.5%
Florida	6,791.00	20976.812	1609.0	3063.834	23.69%	14.61%	8:15	26.4%
New Jersey	7,171.00	8888.543	1797.0	1442.509	25.06%	16.23%	8:18	23.5%
New York	18,241.00	19590.719	4358.0	2954.768	23.89%	15%	8:15	27%
Virginia	4,651.00	8465.207	1197.0	1358.561	25.74%	16.05%	8:23	17.8%
Wisconsin	4,418.00	5792.051	1203.0	947.144	27.23%	16.35%	8:06	28.9%

	Total Population (thousands)		Students (thousands)		Students as share of total population (%)		Average start time	Percentage starting before 8 a.m.
	1970	2018	1970	2018	1970	2018		
Arizona	1,775	7,049	486	1,203	27.4	17.0	8:14	29.5
Florida	6,791	20,977	1,609	3,064	23.7	14.6	8:15	26.4
New Jersey	7,171	8,889	1,797	1,443	25.1	16.2	8:18	23.5
New York	18,241	19,591	4,358	2,955	23.9	15.0	8:15	27.0
Virginia	4,651	8,465	1,197	1,359	25.7	16.0	8:23	17.8
Wisconsin	4,418	5,792	1,203	947	27.2	16.4	8:06	28.9

Fig. 10. An example of redesigning a table from one where the headers are not differentiated from the body text, text and numbers are not properly aligned, and the number of digits is inconsistent.

- Numbers are centered rather than right-aligned, which makes them more difficult to read;
- Likewise, the text in the first column is centered rather than left-aligned;
- Numbers in the table have varying numbers of decimals, which throws off the alignment and suggests a level of precision that is not necessarily true or useful;
- Inconsistent presentation of "percent" labels as text and percentage symbols.

We can easily remedy all these issues with some purposeful editing (bottom of Fig. 10), all done in Excel without too much time or effort:

- Differentiate the column headers and column spanners (labels that cross two columns) with boldface and lines;
- Text and numbers are now properly aligned so that it is much easier to see the largest and smallest values;
- Numbers are right-aligned and centered under the column headers;
- The number of decimals is reduced and made consistent across the entire table; and
- A small bar chart is added to the last column, here simply to demonstrate that adding bars, colors, icons, or other visual elements can be inserted into a table to make patterns clearer and easier to find and read.

Check Your Watch: Yet Another Way to Visualize Time Data

As I reach the end of this tour through some of the many data visualizations available to even the casual researcher and data analyst—again, all created in Microsoft Excel—I embrace the idea of *time* in the data. Instead of a map, bar chart, or other graph, I utilize the actual concept of time as a basis for visualizing the data and place the values around the face of a clock (Fig. 11). Here, state abbreviations are positioned around a clockface at the minute when schools start. Color is used to differentiate states where school starts before the 8:00 a.m. hour.

You might think of this as a fun visualization—I certainly did when I created it in Excel using a series of scatterplots (Schwabish, 2021b)—but it also more clearly communicates the precise *exact* starting times than some of the previous graphs. Users can find their state by searching around the clockface and the shape directly invokes the concept of time.

What would the casual user of the NCES data think if they came to the webpage and saw this big clockface and a "Download the Data" button off to the side? Would they be more likely to engage with it? Would they be more likely to look for their state than scan through a dense table? Again, which graphic form you choose to use will depend on your goals and the needs and expertise of your reader, user, and audience member.

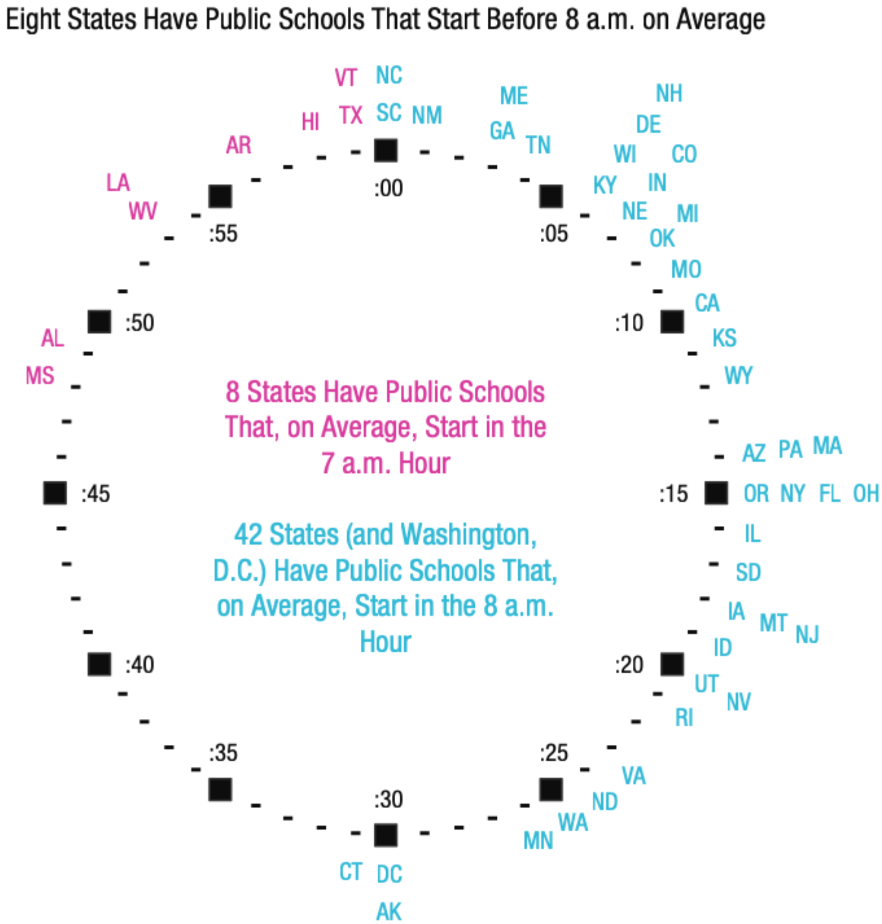


Fig. 11. Another alternative to showing the time-of-day data by embracing the concept of time and placing the data around a clockface. Data source: National Center for Education Statistics (2018).

The Curtain Closes: Conclusion

In my experience, many researchers get hung up on thinking they need to be an expert graphic designer or computer programmer to create great, effective, engaging data visualizations. It is closer to the truth that *anyone* can create such visualizations if they think first about their audience and learn a few, basic strategies and techniques to better data communication.

It also does not take fancy or expensive tools to create effective visuals. All of the graphs presented here were made using Microsoft Excel (see Schwabish, 2021d). And although advanced data visualization tools—for example, programming languages such as R, JavaScript, and Python—have numerous advantages over traditional tools such as Excel, effective visualizations can still be created in what many might view as "basic" tools.

With the increased flexibility of even basic software programs, and a wide range of resources, books, and blogs about data visualization and data communication, researchers and analysts can more effectively communicate their work to help people find patterns, elicit insight, and make discoveries.

Transparency

Editor: Nora S. Newcombe

Declaration of Conflicting Interests

The author receives royalties from Columbia University Press for publications on this topic. He declared that there were no other potential conflicts of interest with respect to the authorship or the publication of this article.

References

Allen, M. (2018, June 16). Why care costs so much. *Axios*. https://www.axios.com/newsletters/axios-am-21c84f86-2705-47a6-a288-b94783d117fb.html?chunk=5#story5

Boy, J., Pandey, A. V., Emerson, J., Satterthwaite, M., Nov, O., & Bertini, E. (2017). Showing people behind data: Does anthropomorphizing visualizations elicit more empathy for human rights data? In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, by Association for Computing Machinery* (pp. 5462–5474). Association for Computing Machinery.

Cesal, A. (2020, July 13). How to create brand colors for data visualization style guides: Your brand colors don't work for data visualization. *Medium*. https://medium.com/nightingale/how-to-create-brand-colors-for-data-visualization-style-guidelines-dbd69c586dd9

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, *79*(387), 531–554.

Cleveland, W. S., & McGill, R. (1985). Graphical perception and graphical methods for analyzing scientific data. *Science*, *229*(4716), 828–833.

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press.

Dornauer, M. E., & Bryce, R. (2020, October 28). Too many rural Americans are living in the digital dark. The problem demands a new deal solution. *Health Affairs Blog*. https://www.healthaffairs.org/do/10.1377/hblog20201026.515764/full/

Engel, P. (2014, February 18). Gun deaths in Florida: Number of murders committed using firearms. *Business Insider*. https://www.businessinsider.com/gun-deaths-in-florida-increased-with-stand-your-ground-2014-2

Few, S. (2004). *Show me the numbers*. Analytics Press.

Field, K. (2018). *Cartography: A compendium of design thinking for mapmakers*. Esri Press.

Field, K. (2021). *Thematic mapping: This is my truth tell me yours*. Esri Press.

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. https://doi.org/10.1177/15291006211051956

Gelman, A. (2011). Why tables are really much better than graphs. *Journal of Computational and Graphical Statistics*, *20*(1), 3–7. https://doi.org/10.1198/jcgs.2011.09166

Groeger, L. V. (2014, September 25). A big article about wee things. *ProPublica*. https://www.propublica.org/nerds/a-big-article-about-wee-things

Hagen, R., & Golombisky, K. (2013). *White space is not your enemy: A beginner's guide to communicating visually through graphic, web & multimedia design*. Routledge.

Heer, J., Bostock, M., & Ogievetsky, V. (2010). A tour through the visualization zoo. *Communications of the ACM*, *53*(6), 59–67. https://doi.org/10.1145/1743546.1743567

Hendel, R. J. (2021, June). Enhancing the pedagogical utility of tables in actuarial teaching. *Expanding Horizons*. https://www.soa.org/sections/education-research/educ-research-newsletter/2021/june/ehn-2021-06-hendel/

Kastellec, J. P., & Leoni, E. L. (2007). Using graphs instead of tables in political science. *Perspectives on Politics*, *5*(4), 755–771. https://doi.org/10.1017/S1537592707072209

Lallanilla, M. (2014, April 23). Misleading gun-death chart draws fire. *LiveScience*. https://www.livescience.com/45083-misleading-gun-death-chart.html

Le Gassik, L., Elavsky, F., & Moritz, D. (n.d.). *Dataviz accessibility resources*. https://github.com/dataviza11y/resources

Lupi, G. (2017, January 30). Data humanism: The revolutionary future of data visualization. *PrintMag*. https://www.printmag.com/article/data-humanism-future-of-data-visualization/.

Morais, L., Jansen, Y., Andrade, N., & Dragicevic, P. (2020). Showing data about people: A design space of anthropographics. *IEEE Transactions on Visualization and Computer Graphics*. Advance online publication. https://doi.org/10.1109/TVCG.2020.3023013

National Center for Education Statistics. (2018). Average public school start time and percentage distribution of school start time, by state: 2017–18. National Teacher and Principal Survey, U.S. Department of Education. https://nces.ed.gov/surveys/ntps/tables/ntps1718_table_05_s1s.asp

Olson, J. M., & Brewer, C. A. (1997). An evaluation of color selections to accommodate map users with color-vision impairments. *Annals of the Association of American Geographers*, *87*(1), 103–134. https://doi.org/10.1111/0004-5608.00043

Scarr, S. (2011, December 17). Iraq's bloody toll. *South China Morning Post*. https://www.scmp.com/infographics/article/1284683/iraqs-bloody-toll

Scarr, S. (n.d.). *Iraq's bloody toll*. http://www.simonscarr.com/iraqs-bloody-toll

Schwabish, J. A. (2016, November 30). Style guide collection. *PolicyViz*. https://policyviz.com/2016/11/30/style-guides/

Schwabish, J. A. (2020). Ten guidelines for better tables. *Journal of Benefit-Cost Analysis*, *11*(2), 151–178. https://doi.org/10.1017/bca.2020.11

Schwabish, J. A. (2021a). *Better data visualizations: A guide for scholars, researchers, and wonks*. Columbia University Press.

Schwabish, J. A. (2021b, May 11). Four (plus) ways to visualize geographic time data. *PolicyViz Blog*. https://policyviz.com/2021/05/11/fourplus-ways-to-visualize-geographic-time-data/

Schwabish, J. A. (2021c, April 19). Further thoughts on developing a style guide at your organization. *PolicyViz Blog*. https://policyviz.com/2021/04/19/further-thoughts-on-developing-a-style-guide-at-your-organization/

Schwabish, J. A. (2021d). A guide to advanced data visualization in Excel 2016/Office365. *PolicyViz*. https://policyviz.com/product/a-guide-to-advanced-data-visualization-in-excel-/2016

Schwabish, J. A. (2021e, March 16). Why your organization needs a data visualization style guide. *PolicyViz Blog*. https://policyviz.com/2021/03/16/why-your-organization-needs-a-data-visualization-style-guide/

Schwabish, J. A., & 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

Tait, A. R., Voepel-Lewis, T., Zikmund-Fisher, B. J., & Fagerlin, A. (2010). The effect of format on parents' understanding of the risks and benefits of clinical research: A comparison between text, tables, and graphics. *Journal of Health Communication*, *15*(5), 487–501. https://doi.org/10.1080/10810730.2010.492560

Tufte, E. R. (1983). *The visual display of quantitative information*. Graphics Press.

U.S. Bureau of Labor Statistics. (n.d.). *Unemployment rate, unadjusted*. https://data.bls.gov/timeseries/LNU04000000?years_option=all_years&periods_option=specific_periods&periods=Annual+Data

Wilson, F. (2018, November 14). Why don't we understand statistics? Fixed mindsets may be to blame. *Frontiers Science News*. https://blog.frontiersin.org/2018/11/14/mathematics-statistics-education/