Visualizing the Virus

Judith Canner
November 10, 2020

Judith Canner is Professor of Statistics at California State University, Monterey Bay.

The United States is finally experiencing some recovery after living through the grip of the COVID-19 pandemic, thanks in large part to the development of highly effective vaccines. Other nations and areas with low vaccination rates continue to struggle. Since March 2020, we have been inundated with data, charts, tables, and dashboards all showing information on the spread and impact of COVID-19 locally, nationally, and internationally. My hope is to present some practical information to help you responsibly consume and digest the information you have seen, and will continue to see, about the COVID-19 pandemic. The ideas and principles presented here may be transferred to understanding visualizations of elections, racial injustice, access to education, or really any topic that is too complex for a single graph to encapsulate.

I do want to take a moment, though, to acknowledge both the individual and collective grief we have all experienced this past year—whether we lost a loved one, missed our families, or struggled with employment. I will be presenting visualizations that attempt to capture the impact of that devastation and loss and will strive to do so with compassion, even when discussing the technical aspects of the visualization itself.

DATA-GENERATION PROCESSES

We cannot talk about visualizations without talking about the source of the data behind the visualization. The scientific process is often viewed as the ultimate source of trustworthy data, but we often do not have a full understanding of how the science actually functions. First, we need to remember the scientific process is slow and that change is a natural part of science.[1] In most situations, we do not have to think about all the research and time it took to develop a new heart medication. We experience the end result only. Science, though, as a whole, moves very slowly, and failure is part of the process—a very important part of the process. Consider the stories of the failures of Thomas Edison before he determined the best material for the light bulb. In the middle of a global pandemic, there is no time to allow for failure as every decision has life-altering effects.

Second, it is also important to recognize that there are many different types of data-generating processes. A case study works at a small scale and uses observation, while a randomized clinical trial may be much larger in scale and experimental so that we may evaluate causal associations. These processes provide different hierarchies of evidence, and, while some evidence is better than others, it all feeds into our understanding of a situation.

Third, we must also consider the source of the information, which is why it is also important to provide my professional credentials. I am a statistician and a biomathematician. I can look at data, I can interpret data, and I can make assessments, but I'm not an epidemiologist, a virologist, or a public health specialist. Many people have touted their expertise on social media, but always be aware that a "Dr." or a "Ph.D." with a name does not make someone an expert in everything. Some of the biggest purveyors of disinformation during the pandemic have come from some of the most "prestigious" universities, and, of course, even trustworthy sources can make mistakes.

Finally, because the purpose of visualization is to explore data, we must remember that data visualizations do not confirm theories or prove causality, but they do allow us to parse out particular patterns and gain understanding. We will focus on three principles that must be considered regardless of context: 1) visualizations must communicate with a purpose; 2) visualizations must use appropriate comparisons; and 3) visualizations must deal with uncertainty.

COMMUNICATING WITH PURPOSE

In the first weeks of the pandemic as experienced in the United States, especially as lockdowns began in many counties and states, the most popular visualization wasn't based on any actual data at all. It was the "flatten the curve" visualization. The curve was not a new invention but has existed for years in the plans and proposals the Centers for Disease Control and Prevention (CDC) had created to mitigate pandemic influenza (see Figure 1 below). Though not based on data, the purpose of the visualization was to show the public why mitigation factors such as lockdowns, physical distance, masks, and hand-sanitization were vital to keeping our health care system from being overwhelmed by cases. It was extremely effective at communicating its purpose, though at the time I do not think we collectively realized just how bad the pandemic would actually become in the United States.

A visualization made for one purpose may be co-opted and used for another purpose, often incorrectly and not for the purpose the creator intended. In a visualization from March 2020, the *Financial Times* emphasized how fast the virus was spreading in different countries (see Figure 2 below). This visualization compared multiple countries, which had seemed to mitigate the spread of the virus, to the United States and other countries in order to demonstrate possible measures the United States and others might consider for their own mitigation strategies. Let's break down the visualization, as it is a

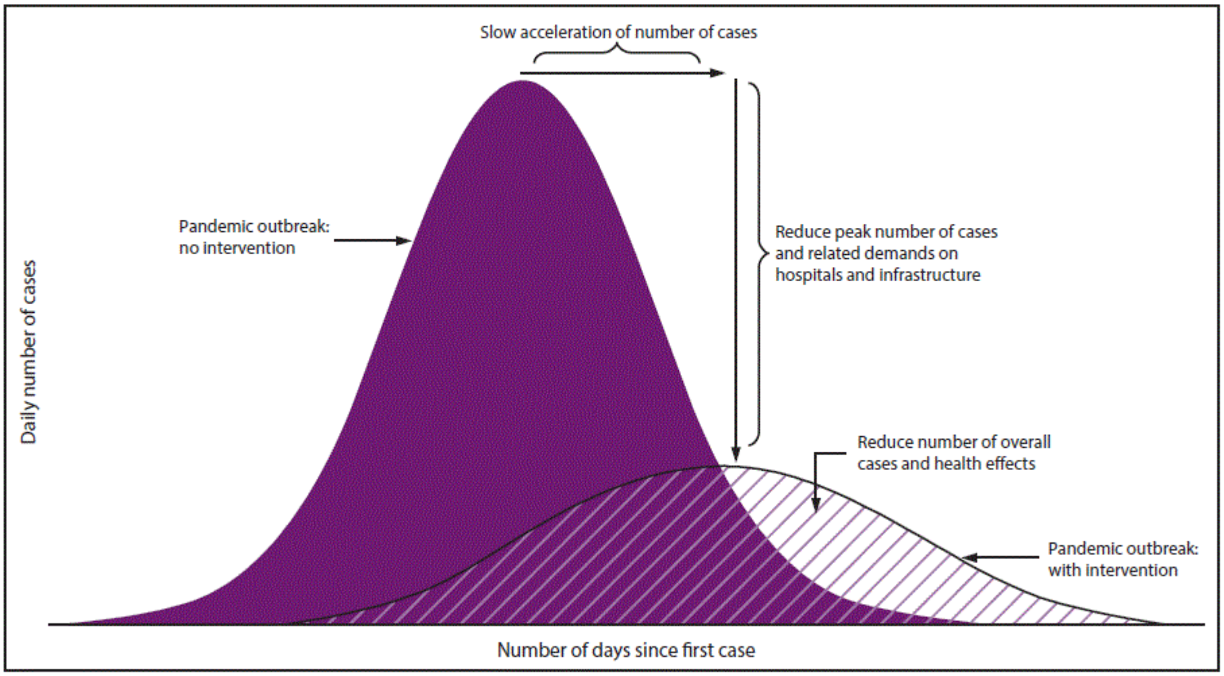

Figure 1. The “flatten the curve” visualization.[2]

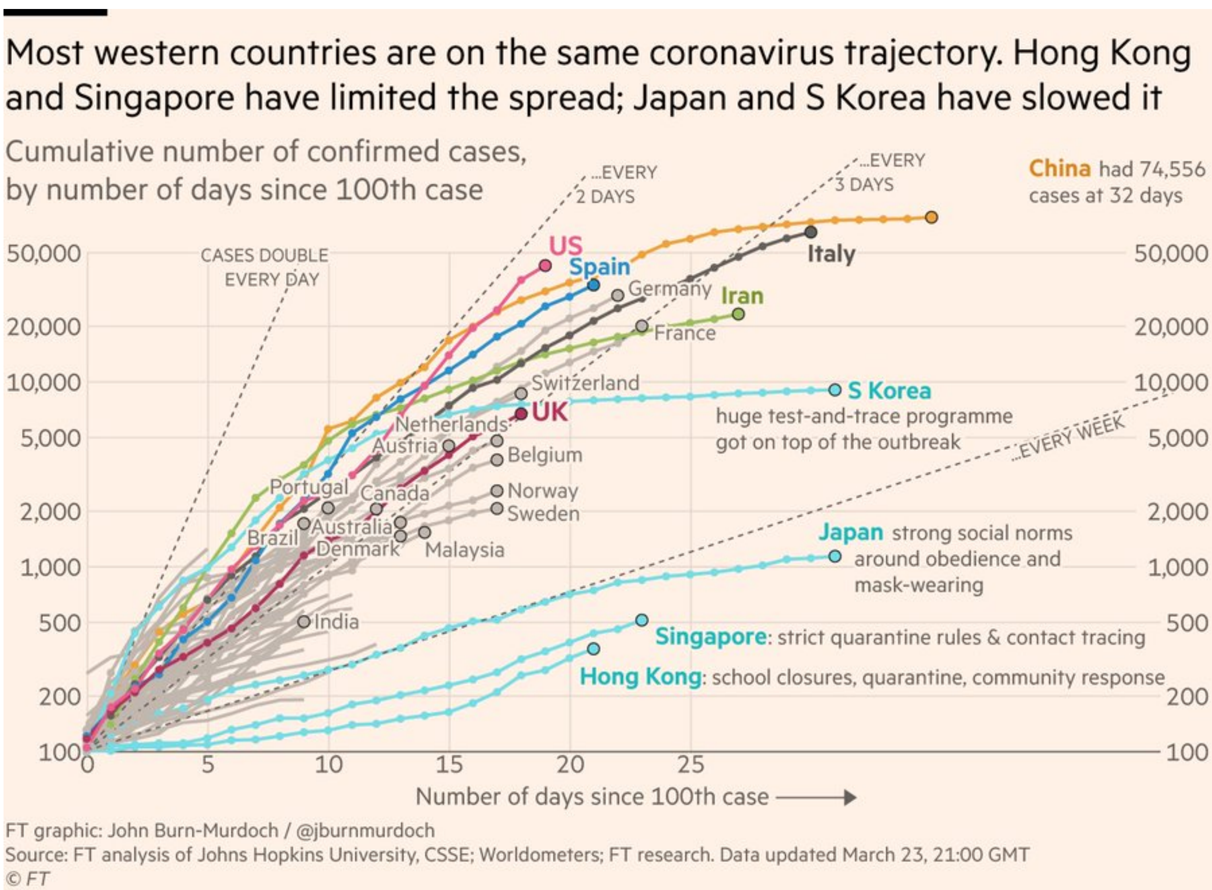

Figure 2. The spread of the virus in different countries since the hundredth case. Figure created by John Burn-Murdoch.[3]

very busy graph:

- There are many lines, each representing a different country. Some are highlighted in color while others are greyed out. Some have labels, other do not.
- There are four countries emphasized with the same teal color for each of their lines, South Korea, Japan, Singapore, and Hong Kong. Each country is also annotated to indicate their chosen mitigation measures.
- The x-axis does not represent a specific date but the number of days since 100 recorded cases. Since the spread of the virus began at a different point in time for each country, scaling the data to the number of days since 100 cases instead of using the raw date allows for a more direct comparison of spread over time.
- The y-axis, which represents the cumulative number of confirmed cases, does not have a linear scale but a logarithmic scale. We will discuss the choice of logarithmic scale in more detail, but, for now, notice that the scale allows us to compare the rate of change in the spread instead of overall differences.
- Additional lines were added that don't represent specific countries but demonstrate what the trajectory would look like for a country where cases doubled over a specific period of time (e.g., every week, every two days, every day).
- Finally, the title emphasizes the creator's intent—to demonstrate that most western countries had a similar trajectory, but the four highlighted countries were on a different trajectory.

The creator's intention becomes clear: if Hong Kong and Singapore have limited the spread and Japan and South Korea have slowed it, perhaps this is due to their mitigation measures. Recall that the visualization was published just as various locations in the United States were beginning lockdowns, masks, and other mitigation measures. Soon, though, such "spaghetti plots" began to be used to justify the actions, or lack of actions, of the United States. "We're doing better than this or that country!" became the exclamation, almost as if we were in a strange race. The use of the visualization no longer remained in the control of the creator.[4] Of course, other visualizations quickly began to appear, seemingly based on data and with seemingly contradictory information about the potential deadliness of the SARS-COV-2 virus and its spread.[5] In the end, though, the comparisons made in many visualizations, though perhaps based on real data, were not appropriate comparisons.

MAKING APPROPRIATE COMPARISONS

The purpose of the spaghetti plot was to make a comparison to suggest that certain mitigation measures had the potential to be effective in the United States. It may seem reasonable to make comparisons of cumulative counts at the earliest time point of spread, and the creator of Figure 2 tried to account for many differences in the manifestation of the pandemic in different locations. The decisions were appropriate for the purpose. As the pandemic progressed, however, such comparisons became less appropriate. Each country collected its data in different ways, using different methods and using different tests to confirm infections. Even if we could control for all these sources of variability in our data on confirmed cases, we must still take care when drawing comparisons between different populations or groups at different scales.

Ecological Fallacies

Let us first consider scale. While by definition a pandemic is worldwide, it also manifests locally in different ways. One location may be affected differently for a variety of reasons, and patterns at a national or state level may differ from patterns at a county or city level. Therefore, we must consider ecological fallacies, the idea that the individual behaves the same way as the whole. Researchers frequently try to infer the behavior of individuals based on the behavior of the whole. For example, we can observe the United States during the summer of 2020 to see the difference of national-level trends versus state-level trends in deaths due to COVID-19 (see Figure 3). The United States may be one nation,

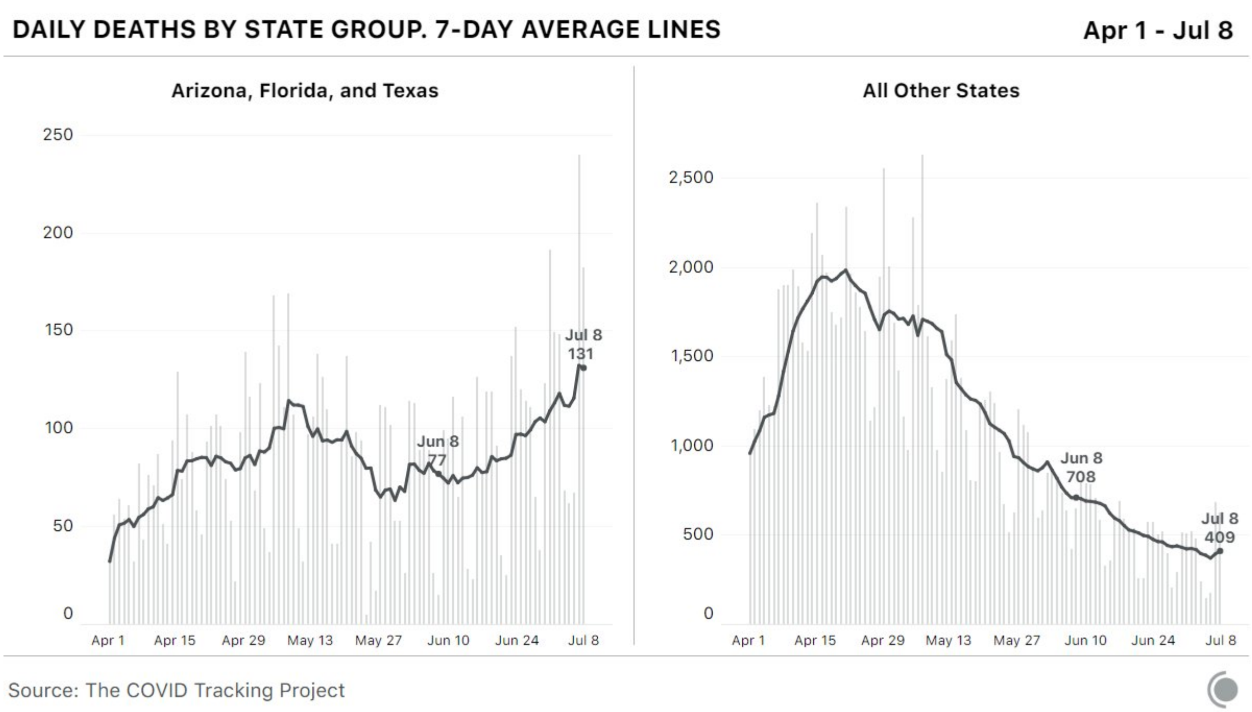

Figure 3. National-level trends vs. state-level trends in deaths due to COVID-19. The left graph represents daily deaths for Arizona, Florida, and Texas combined, and the right graph represents daily deaths for all other states from April 1, 2020, to July 8, 2020. Image credit: The COVID Tracking Project, July 9, 2020, https://covidtracking.com/.

but it is made up of fifty different states plus territories, each with different mitigation measures. The national trend seemed to demonstrate improvement, as cases, hospitalizations, and deaths all appeared to decrease. But if we examine the three states that at the time had some of the highest case rates—Arizona, Florida, and Texas, we see that deaths were increasing across the three states in contrast to the national trend. The observations of different trends at the national and state level are a type of ecological fallacy called an amalgamation paradox or Simpson's paradox, where a trend disappears or reverses when you aggregate the data by subgroups. Of course, we eventually saw the national level cases, hospitalizations, and deaths increase to the horrific counts that there were in winter 2020-2021, but, even then, not all locations experienced the pandemic in the same way.

Logarithms and Scales

The other issue of scale is the size of the groups being compared. We saw this in our spaghetti plot (Figure 2 above), where the comparison of different countries can often obfuscate the tremendous differences between them. First, though, let us consider the scale of the axis used in the visualizations, as the choice of axis scale often causes confusion. Using a linear scale for counts of daily confirmed cases, we can see how the pandemic has progressed over time (see Figure 4a below), but the spread of a virus is a non-linear process. If someone is infected, they do not infect just one other person but many others. Therefore, because the spread of COVID-19 is an exponential process, the rate of spread is better represented with the count data mapped to a logarithmic scale (see Figure 4b below). Different scales are not meant to be deceptive, for both scales are useful, but they communicate different aspects of the data. Of course, comparing case totals across three countries that have vastly different population sizes may also encourage inappropriate comparisons by the audience, even if unintended by the creator (see Figures 5a and 5b below for a comparison of Figures 4a and 4b scaled by population size).

Maps and Scales

One of the first visualizations that everyone looked at were maps generated by the Johns Hopkins University COVID-19 Dashboard. Maps also played a large role in many early conspiracy theories about the pandemic. For example, at one point a viral post showed a map of COVID-19 cases and the density of 5G cell service towers in the United States, insinuating a connection.[6] However, if we simply looked at a map of the population density of the United States, it would appear to have the same "connection" because COVID-19 cases and 5G tower density varies with population density (we call this a spurious correlation*)*.

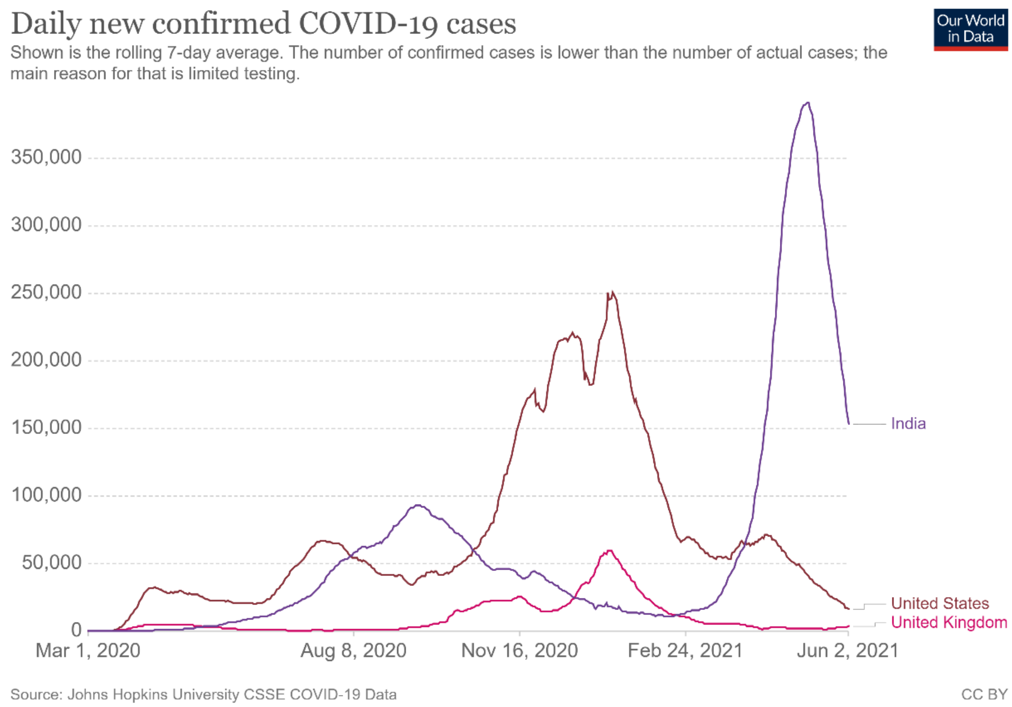

Figure 4a. Comparison of daily new confirmed COVID-19 cases in the United States, the United Kingdom, and India since March 1, 2020 (with counts represented on a linear y-axis). Image credit: "Coronavirus (COVID-19) Cases," *Our World in Data*, https://ourworldindata.org/covid-cases.

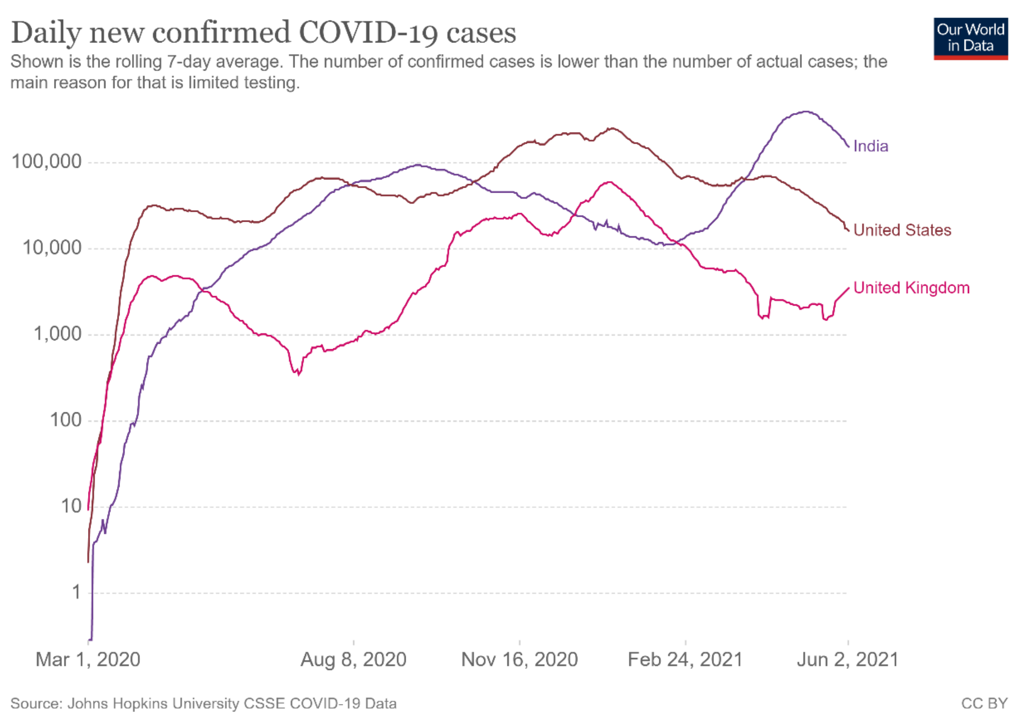

Figure 4b. Comparison of daily new confirmed COVID-19 cases in the United States, the United Kingdom, and India since March 1, 2020 (with counts represented on a logarithmic (base 10) y-axis). Image credit: "Coronavirus (COVID-19) Cases," *Our World in Data*, https://ourworldindata.org/covid-cases.

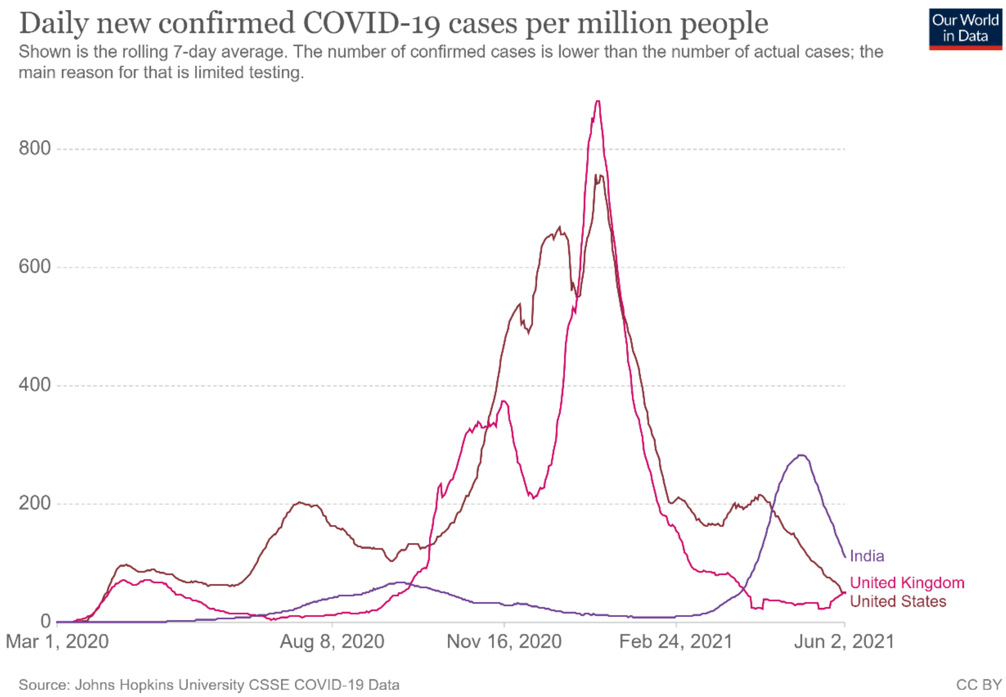

Figure 5a. Comparison of daily new confirmed COVID-19 cases per one million people in the United States, the United Kingdom, and India since March 1, 2020 (with counts represented on a linear y-axis). Image credit: "Coronavirus (COVID-19) Cases," *Our World in Data*, https://ourworldindata.org/covid-cases.

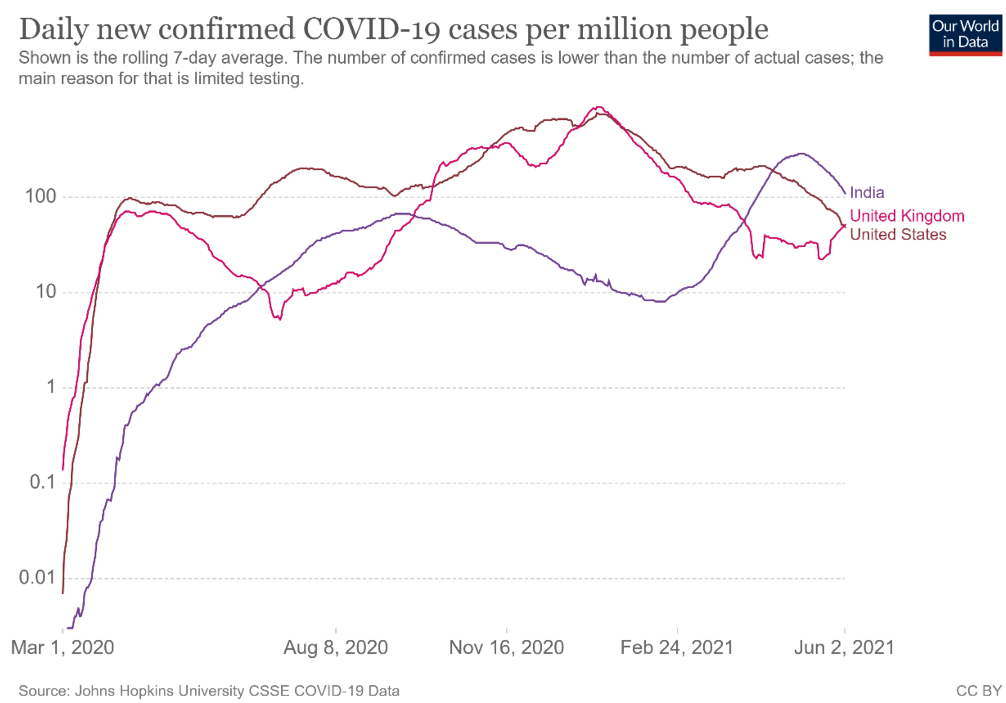

Figure 5b. Comparison of daily new confirmed COVID-19 cases per one million people in the United States, the United Kingdom, and India since March 1, 2020 (with counts represented on a logarithmic (base 10) y-axis). Image credit: "Coronavirus (COVID-19) Cases," *Our World in Data*, https://ourworldindata.org/covid-cases.

To account for population variability among the locations we compare, we can consider reporting rates rather than counts (although both are used for different purposes). The Coronavirus dashboard from National Public Radio (e.g., see Figure 6) allows the user to select for either cases or deaths by total

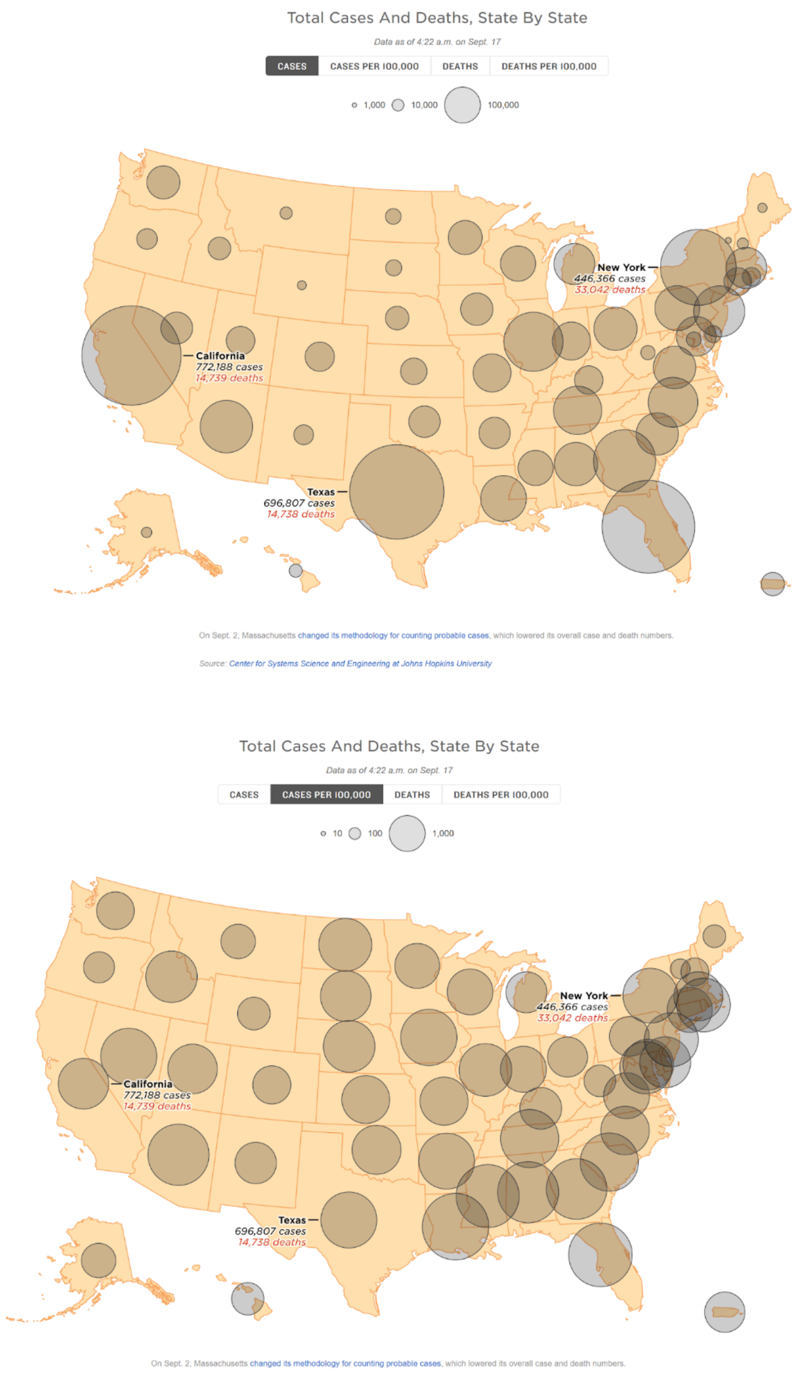

Figure 6. Map of the United States with total cases by state (top) and with total cases per 100,000 by state (bottom), from September 17, 2020. Image credit: National Public Radio, [https://www.npr.org/sections/health-shots/2020/09/01/816707182/map-tracking-the-spread-of-the-coronavirus-in-the-u-s.](https://www.npr.org/sections/health-shots/2020/09/01/816707182/map-tracking-the-spread-of-the-coronavirus-in-the-u-s)

counts or by count per 100,000 people to account for the highly variable state population sizes. The bigger the circle, the more cases there are. The creators have annotated a few specific states—New York, California, and Texas—as these states have large case counts and, therefore, larger circles, but when we visualize the count scaled by population size, or cases per 100,000, those circles start to look a little more similar (at the date of observation). The states with the largest population sizes will also have the largest case counts. Each graph communicates something different. The top graph in Figure 6 emphasizes the total impact on human life, the counts. The bottom graph in Figure 6 emphasizes the expanse of each outbreak at the state level, the cases per 100,000.

Another common visualization type used to display case data are choropleth maps, which use color scales to represent counts or rates. In the summer of 2020, two maps were posted in the Georgia Department of Public Health Daily Status Report fifteen days apart representing cases per 100,000 people within each county (see Figure 7). If you just look at them without reading the scales, it appears that not much had changed in the case rates, but the two graphs have very different scales for each color on the map. Figure 8 below shows that if we rescale the July 17 map in Figure 7, right, to the same scale as the July 2 map in Figure 7, left, the reality that cases increased over the two weeks becomes clearer.

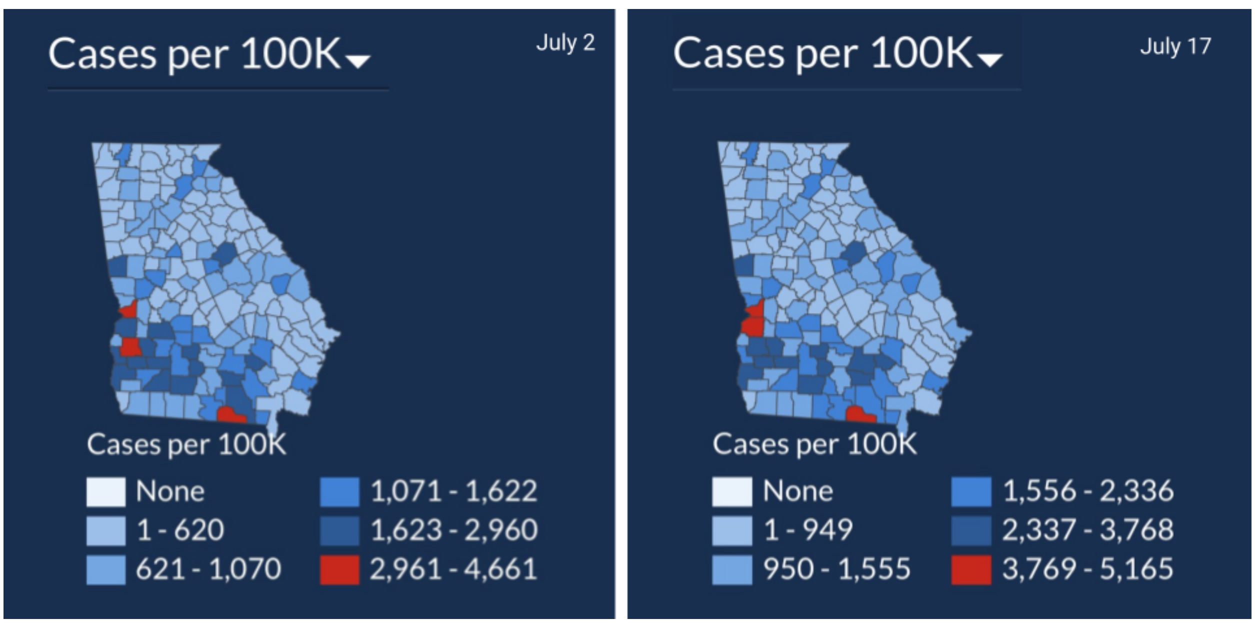

Figure 7. Screenshots of maps from the Georgia Department of Public Health Daily Status Report (https://dph.georgia.gov/covid-19-daily-status-report) from July 2, 2020 (left) and July 17, 2020 (right). Notice the differing color scales for each map.

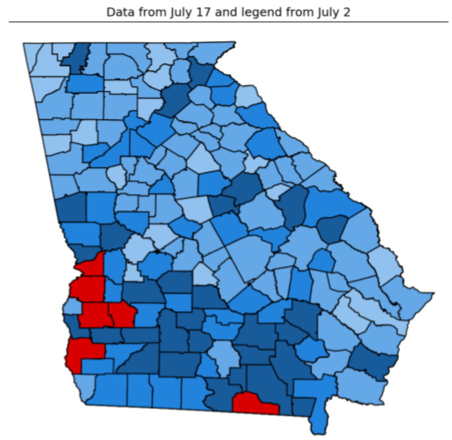

Figure 8. Screenshot of a rescaled map, colored using the scale from July 2, 2020. Image credit: A now-deleted tweet by Andisheh Nouraee @andishehnouraee.

Although we might want to jump to accusations of malice, remember that, at this point in the pandemic, many states were still developing their COVID-19 data dashboards (notice the difference between the screenshots in Figure 7 and the current dashboard). Data visualization is a skill that takes time to develop, and, as we have already established, time was in short supply. In addition, many visualization software programs, for example Tableau, provide automatically chosen scales based on the spread of the data. A trained creator will know to consider the scale and how to adjust it, but whoever was in charge of the Georgia dashboard at this point in the pandemic may not have had this training. Finally, we must also realize that maps are not usually meant to be viewed over time. They are meant to capture the data at a specific time point.

Furthermore, the granularity of a choropleth can make a big difference in comparisons. Consider a visualization of the percentage of people wearing masks in public most or all the time (see Figure 9 below) by state. The graph largely reflects mask mandates at the time, but it does not necessarily provide a complete picture. Consider instead a map of the chance that all people are wearing masks in five random encounters (see Figure 10 below).[7] The county-level representation provides a very different perspective on how mask-wearing varied within each state (although we must consider that Figures 9 and 10 are from different time points). Maps can provide great geographic comparisons, but we must remember an important truth when using them to draw conclusions: land does not get COVID-19 or wear masks; people get COVID-19 or wear masks. We must maintain awareness of the potential for making an ecological fallacy, especially with maps.

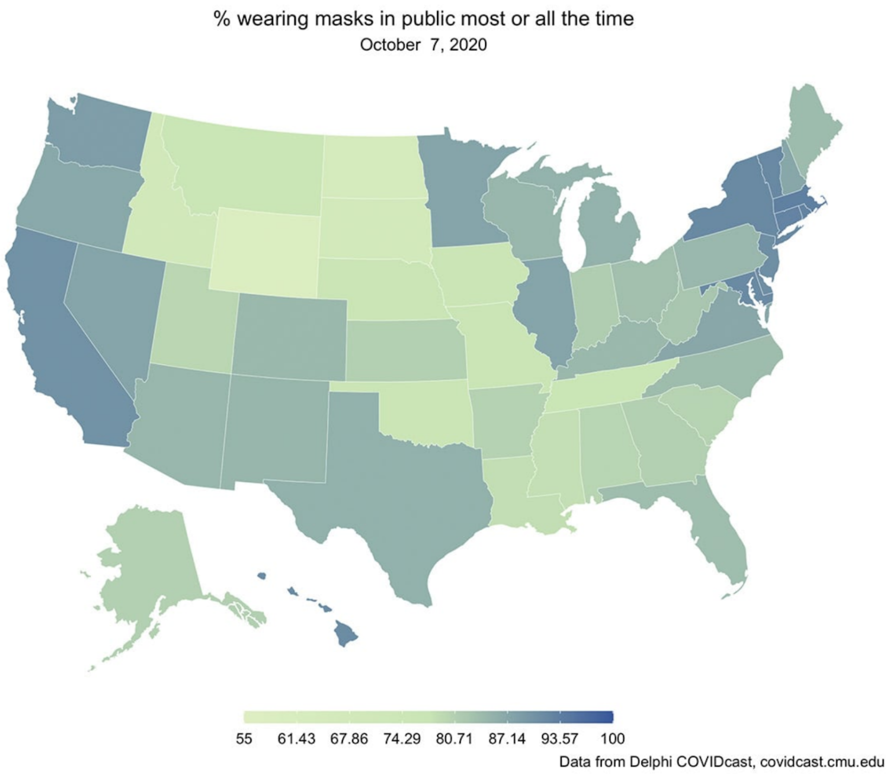

Figure 9. Percentage of people wearing masks in public most or all the time. Image credit: The Carnegie Mellon University's report *COVIDcast Now Monitoring Daily U.S. Mask Use, COVID-19 Testing* from October 12, 2020, https://www.cmu.edu/news/stories/archives/2020/october/covidcast-mask-use.html

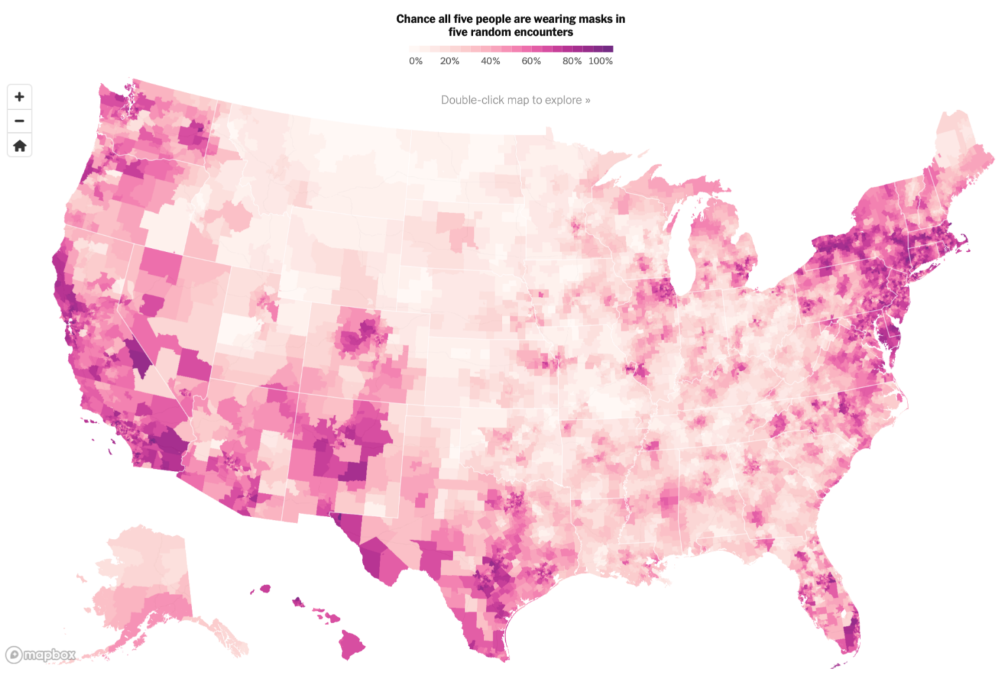

Figure 10. Chance that all five people are wearing masks in five random encounters. Image credit: "A Detailed Map of Who Is Wearing Masks in the U.S.," *The New York Times*, July 17, 2020.

Bad Practices with Scales

We have seen how, even with good intentions, scales can be difficult to manage as we consider their impact on how we communicate our purpose, but sometimes there are just bad practices. When interpreting a visualization, always keep an eye out for misleading axis labels, as in Figure 11 below. If the categories of the x-axis are ordered, such as the dates in Figure 11, they should not be reordered in the visualization because it muddles the viewer's ability to make comparisons between the ordered values

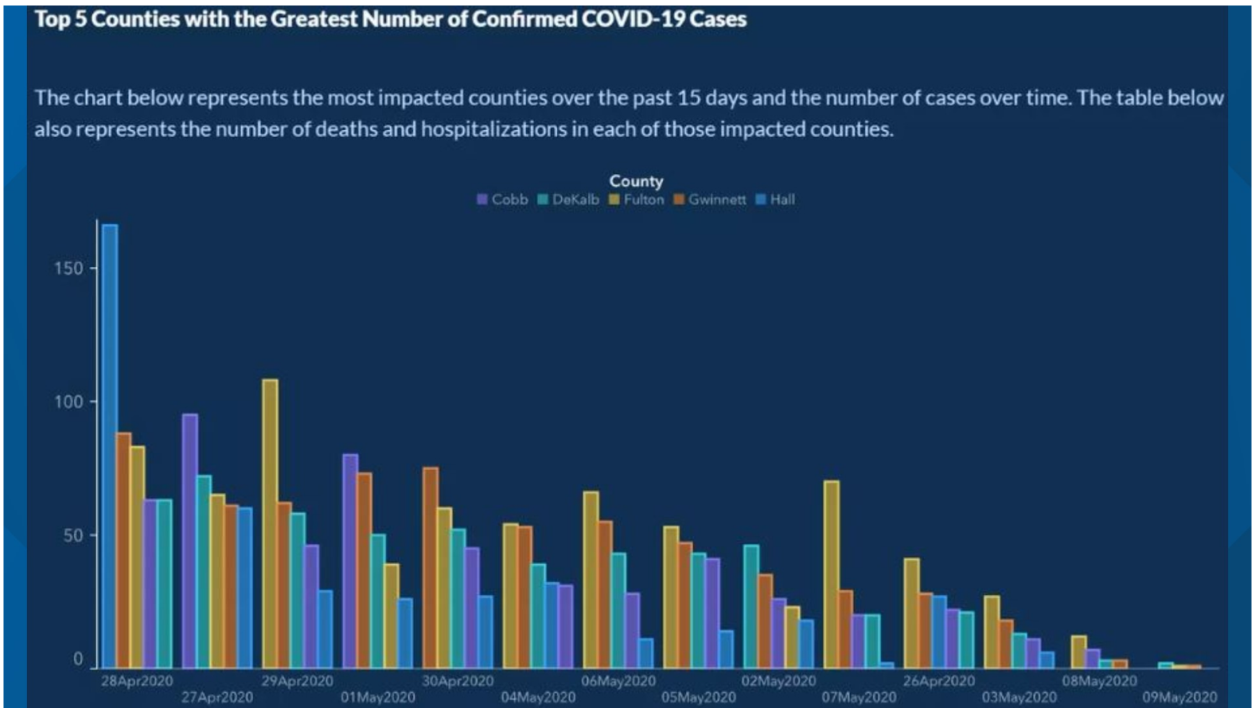

Figure 11. A visualization with misleading axis labels. Image credit: Maps from the Georgia Department of Public Health Daily Status Report, May 2020, https://dph.georgia.gov/covid-19-daily-status-report.

(e.g., from one day to the next day). The truncation of axis scales, such as with the y-axis in Figure 12 below, violates the principle of "proportional ink," which states that if a shaded region represents a numerical value, then the area of the shaded region should be directly proportional to the value itself.[8] Figures 12 and 13 both display inverted y-axis scales, which make trends that are increasing actually look as if they were decreasing. Finally, some visualizations use dual axes. In Figure 14 below, the two different scales for the dual y-axes exaggerate the change in cases per 100,000 in counties that have mask mandates. The dual axes must be scaled proportionately to one another. The y-axis in Figure 14 is also truncated, but since it is a line and not a shaded region that represents a numeric value, the truncation is less likely to confuse us. Again, bad axis scales and types may be the result of uninformed visualization creation, or they may be the result of a malicious intent to mislead. Either way, as consumers of visualizations, we must diligently check that the visualization we are interpreting is appropriately constructed.

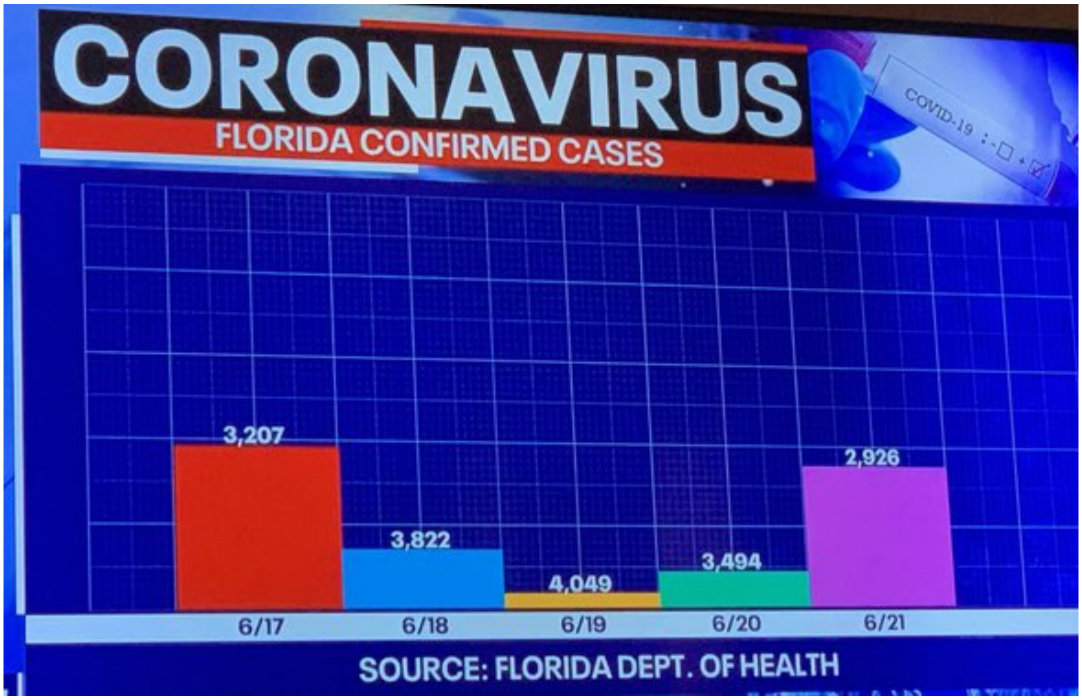

Figure 12. A visualization showing truncated axis scales. Image credit: @DannyPage on Twitter, showing a picture from a local news station in Florida from June 2020.

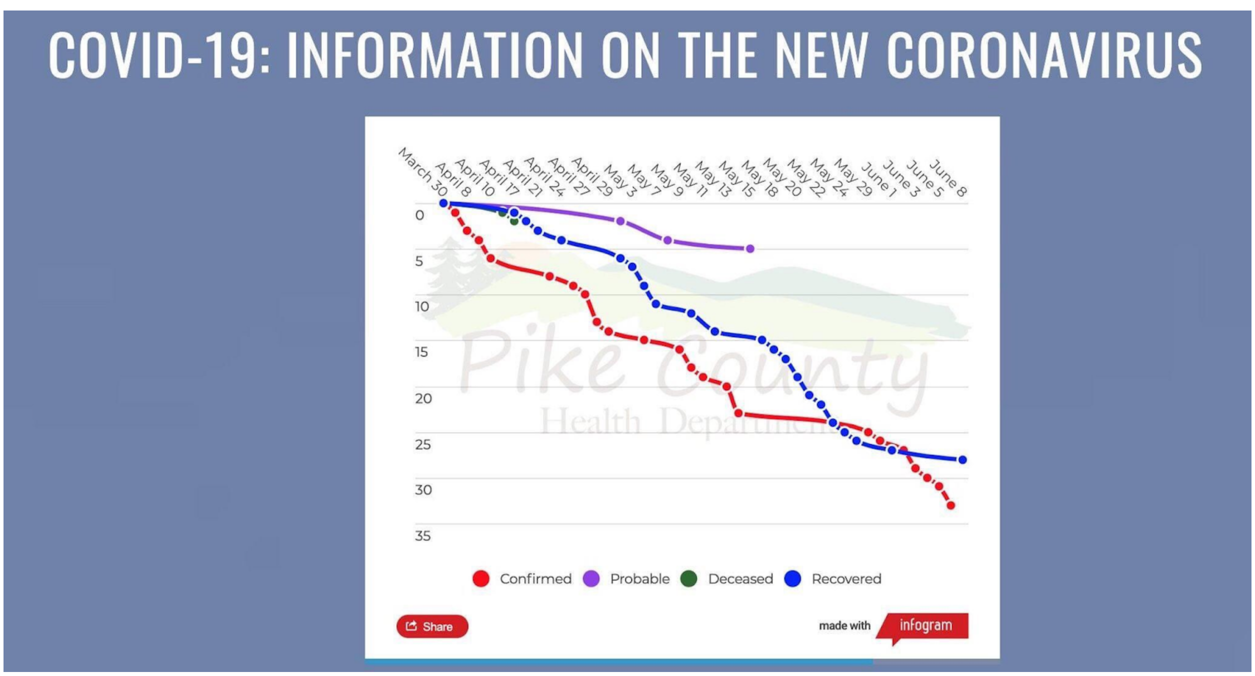

Figure 13. A visualization showing the inversion of axes. Image credit: Pike County Health Department, June 2020 https://www.pikecountyhealth.com/v4i/covid-19.html.

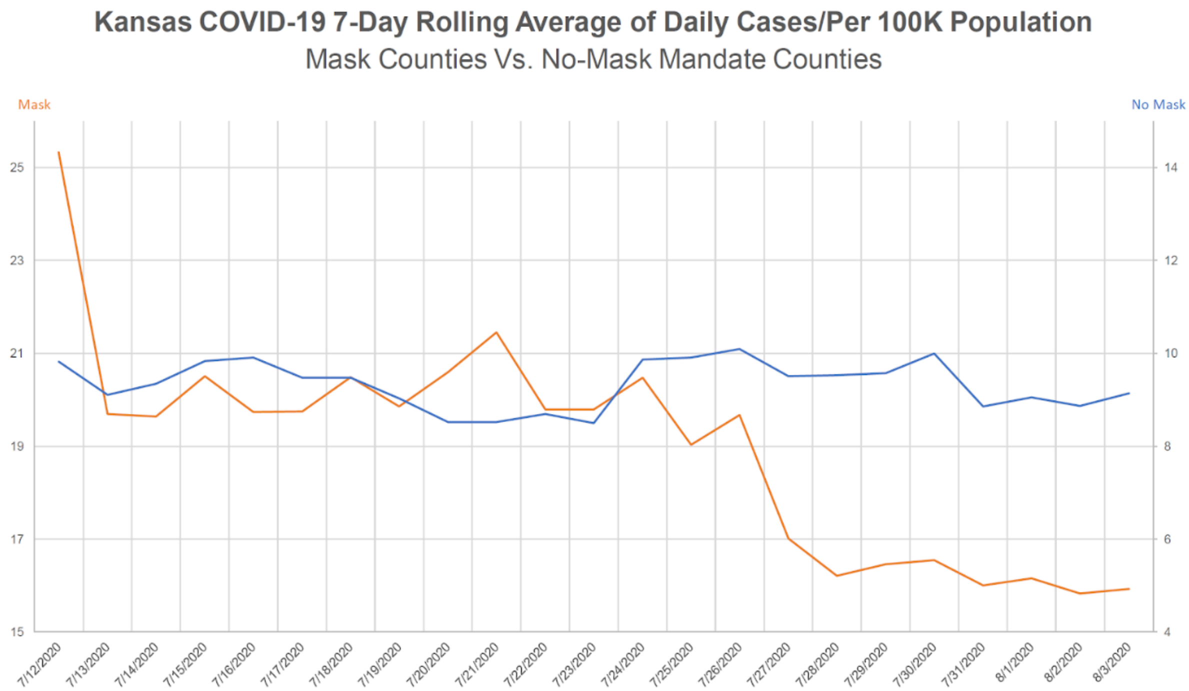

Figure 14. A visualization using dual axes not proportionally scaled to one another. Image credit: A graph shared on the Rachel Maddow Blog, August 2020 (https://twitter.com/MaddowBlog/status/1291553722527604736).

DEALING WITH UNCERTAINTY

On several fronts, uncertainty has been a challenging aspect of living through a pandemic. First, the overall uncertainty due to lack of knowledge about the virus itself and then the uncertainty in mitigation, treatment, and vaccines have all impacted our lives. In the world of visualization, two forms of uncertainty impacted their use for understanding and decision making: 1) uncertainty with data and 2) uncertainty in predictions.

Uncertainty with Data

The uncertainty with data starts with the case counts. This is emphasized through measurement of the case positivity rate, the percentage of all tests that return a positive (COVID-19 infected) result. The CDC, World Health Organization, and other organizations seek a case-positivity rate less than 5% because a low case-positivity rate provides certainty that the numbers of cases recorded are the actual number of cases in a population. All case counts will produce an undercount because it is impossible to test all people, but if a population can maintain a low case-positivity rate, then case counts are more certain.

The other source of uncertainty with data is the lag in reporting test results. Many locations had a two-day (or longer) lag before the individual received their results. That meant reporting to state health officials also had to wait a few days, and even more time elapsed before the numbers became public. We see lags in reporting in the overall structure of the data, as shown in Figure 15, with weekly cycles of few to no case reports on the weekends and then an uptick on Mondays or Tuesdays. To account for such uncertainty, moving averages—usually based on seven-day counts of cases from several days both prior to and after the date represented—are common when visualizing cases, hospitalizations, deaths, and testing over time (see also Figures 3, 4a, 4b, 5a, 5b, and 15). The seven-day average smooths out the variability due to reporting lags to highlight the overall trends in the data.

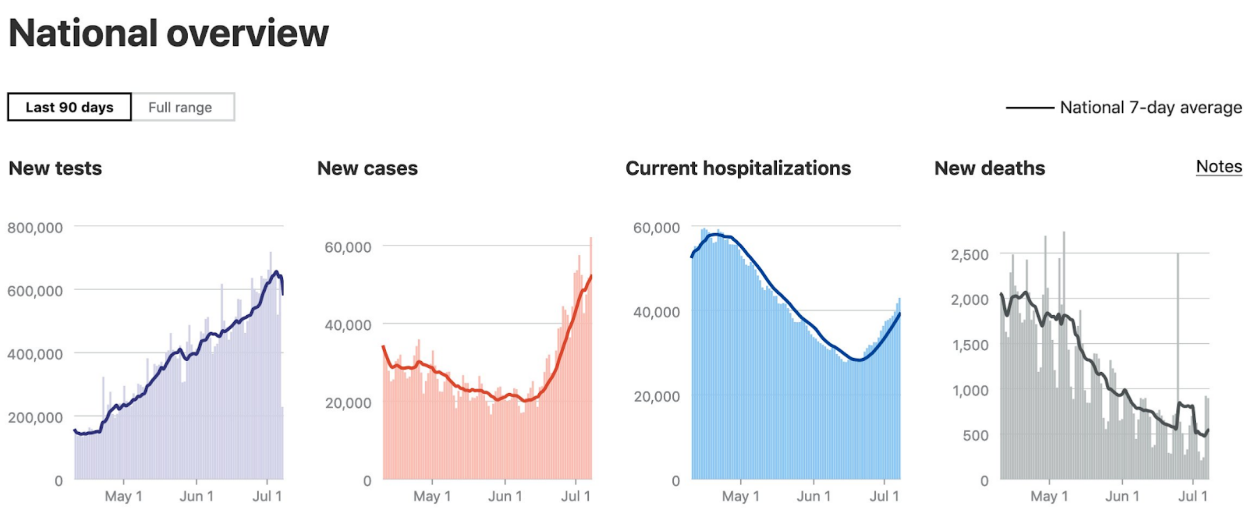

Figure 15. Visualizations that address the uncertainty with reporting test results. These graphs show the national overview of testing, cases, hospitalizations, and deaths. Notice on each graph the darker color line representing the seven-day average overlaid over bars that demonstrate weekly variation. Image credit: The COVID Tracking Project, July 2020.

Uncertainty continues to remain in the actual number of deaths as a result of COVID-19. Data sources are being reviewed and revised regularly to catch redundancies in reporting. Early in the pandemic, the lack of testing led to undercounting of deaths due to COVID-19. Even with better testing, the number of deaths above average shown in Figure 16 below is astounding and heartbreaking to see.

Uncertainty with Predictions

In May 2020, the Council of Economic Advisors shared a visualization that claimed, based on their model, that the number of deaths would go to zero by the end of June 2020 (see Figure 17 below), which led many to think the pandemic would end quickly. There are many issues with the predictions from the visualized models. First, it is dangerous to extrapolate from time series data fit with a mathematical equation (a cubic model) while ignoring the underlying dynamics and complexities of the spread of the virus. The simplest model of disease spread must consider the dynamics of the Susceptible, Infected, and Recovered people in population, or the SIR model. The dynamics of the infection rate,

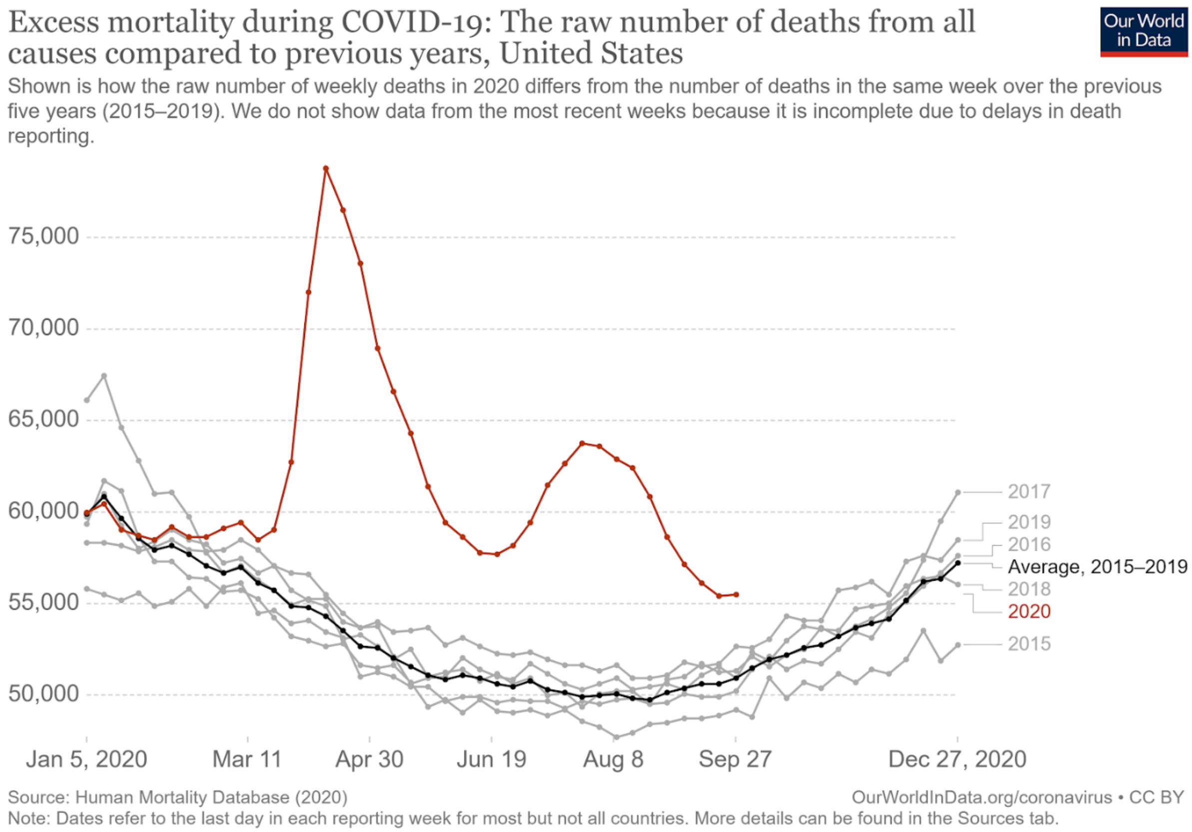

Figure 16. Excess mortality during 2020 in the United States. The red line indicates deaths in 2020 compared to previous years and the average from 2015-2019. Image credit: Our World in Data, https://ourworldindata.org/excess-mortality-covid.

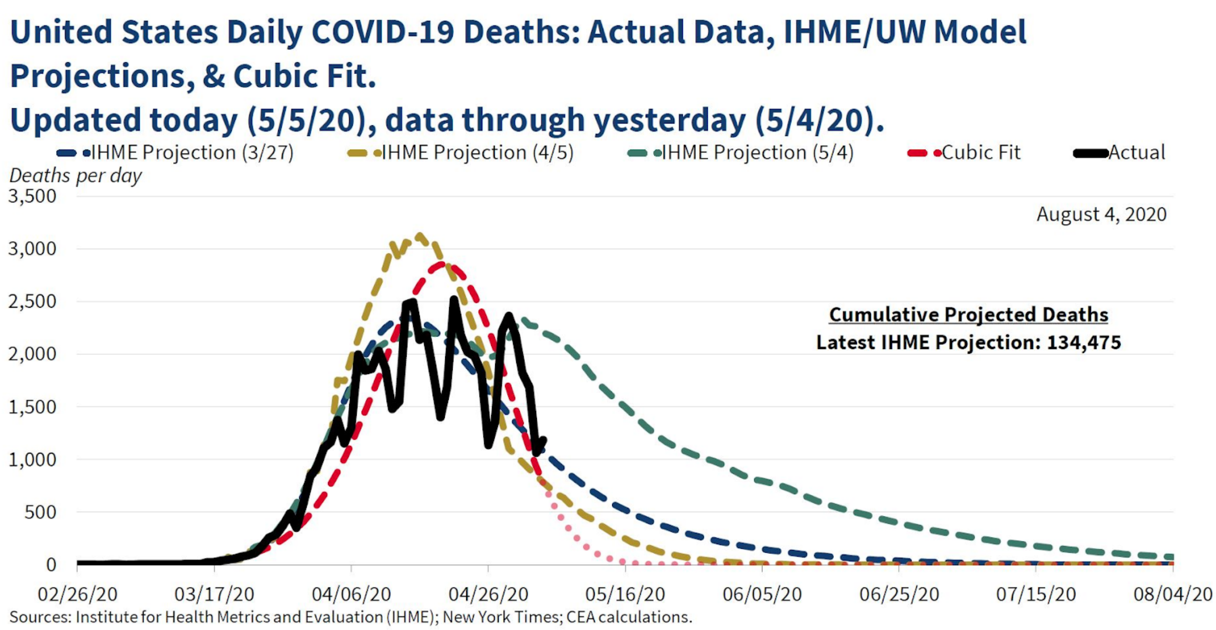

Figure 17. A visualization making a prediction without depicting uncertainty. Image credit: Tweet from the Council of Economic Advisers on May 5, 2020, https://twitter.com/WhiteHouseCEA45/status/1257680258364555264.

incubation period, time to recovery, risk of reinfection, death rates, etc., in addition to considering a connected global population, must be accounted for in the model predictions. The other model from the Institute for Health Metrics and Evaluation (IHME) represented in Figure 17 did consider some of these dynamics, but it was based on a poorly informed understanding of the COVID-19 infection dynamics at the time. The obviously poor projections also lacked the second necessary quality of a prediction: a measure of uncertainty, e.g., confidence intervals or prediction intervals. In practice, any prediction that ignores complex dynamics and does not quantify the uncertainty of the predicted value should be immediately questioned and its veracity examined.

COMPASSION AND VISUALIZATION

We have all been exposed to an overwhelming amount of information, and it is difficult to make sense of it. There is so much information swirling around, not just about COVID-19, and the best first step is to simply pause. We must take the time to evaluate the source of the information and consider how that source would acquire the information presented. Then, before we decide to share a visualization that purports to prove some point, we must consider the visualization's original purpose, ask if it is making appropriate comparisons, and determine if it accounts for uncertainty.

In addition to the technical aspects of visualization, we must also remember to have compassion. Every number represents a person. When the United States passed 100,000 deaths, *The New York Times* created an interactive visualization that emphasized the humanity behind each death (see Figure 18). The United States reached 500,000 deaths in February 2021. *The New York Times* captured the new,

Figure 18. An interactive visualization representing each death from February to May 26, 2020, in the United States. Image credit: *The New York Times*, https://www.nytimes.com/interactive/2020/05/24/us/us-coronavirus-deaths-100000.html.

Figure 19. Front page showing frequency of death during the pandemic. Image credit: *The New York Times,* February 21, 2021, https://static01.nyt.com/images/2021/02/21/nytfrontpage/scan.pdf.


"All the News That's Fit to Print"

The New York Times

Late Edition

VOL. CLXIX ... No. 58,688 — NEW YORK, SATURDAY, MAY 9, 2020 — $3.00

U.S. UNEMPLOYMENT IS WORST SINCE DEPRESSION

MONTHLY CHANGE IN JOBS SINCE THE END OF WORLD WAR II

NATIONAL UNEMPLOYMENT RATE — 14.7%

APRIL UNEMPLOYMENT

White	Black	Hispanic	Men	Women	35 to 44	45 to 54	55+	25 to 34	20 to 24
14.2	16.7	18.9	13.5	16.2	11.5	12.3	13.6	14.5	25.7

20,500,000 jobs lost in April

Georgia Killing Puts Spotlight on a Police Force's Troubled History

Long Path to Arrests of Ex-Officer and Son in Black Man's Death

April's Rate of 14.7% Touches All Parts of Economy

By NELSON D. SCHWARTZ and BEN CASSELMAN

NEWS ANALYSIS

In Flynn Case, Russia Inquiry Is Barr's Target

By MARK MAZZETTI

WHITE HOUSE MEMO

If West Wing Isn't Safe, Is Any Office?

By PETER BAKER and MICHAEL CROWLEY

A Mexico City crematory. Regional virus deaths are officially low.

As Official Toll Ignores Reality, Mexico's Hospitals Are Overrun

By AZAM AHMED

G.O.P. Convention Qualms

Politics and the Justice Dept.

A Walk a Day to Get By

Moscow's Paused Celebration

Banker Shakes Up Suriname

A Good Walk, Unspoiled

Anyone Up to Shoot Hoops?

New York's Return to the Arts

A Shot at Stardom, Revised

Timothy Egan

Figure 20. Front page showing job losses due to the pandemic. Image credit: *New York Times*, May 9, 2020, https://static01.nyt.com/images/2020/05/09/nytfrontpage/scan.pdf.

heartbreaking milestone with a front-page visualization to demonstrate the frequency of deaths over the course for the pandemic (see Figure 19 above). The visualization is profoundly effective in demonstrating the severity of the pandemic over the winter months. However, we no longer see the people but just see the points, and our compassionate response to the data is dulled. Data and compassion are not often paired together. We tend to think of data as some neutral record of the facts, and "facts" do not need compassion to support their meaning. But we must remember to find the right balance between information and compassion, fact and emotion, the statistic and the soul.

One of the most striking visualizations as it relates to the medium in which it was presented appeared in *The New York Times* on May 9, 2020, as shown in Figure 20 below, showing the absolutely staggering loss of jobs in April due to the lockdowns. It is an amazing visualization for its use of the medium (the front page of the newspaper) but also for its emotional impact. The red line plunging down the right-hand edge of the page makes a staggering and heart-rending point. To see the scale of the loss of jobs forces us to think about the impact the pandemic has had—not just for those who have lost their loved ones or who are experiencing the long-term health effects of this virus but also for those who have been devastated economically and are still struggling to survive.

We must continue to be vigilant to understand the lessons learned from the pandemic when it comes to creating, understanding, and communicating through data visualizations because the same lessons apply to many other "big problems" facing our world both now and in the future.

NOTES

1. Pacific Science Center and the University of Washington's Center for an Informed Public, "Facts in the Time of COVID-19," *Genially,* June 17, 2020, https://view.genial.ly/5eea3a0c15e1e60d88c5c4d0/interactive-content-facts-in-the-time-of-covid-19. This site provides a nice overview of things to consider when engaging with new information about the pandemic; its principles can be applied to almost any situation.

2. Centers for *Community Mitigation Guidelines to Prevent Pandemic Influenza—United States, 2017*, but adapted from a 2007 report, http://dx.doi.org/10.15585/mmwr.rr6601a1.

3. This *Financial Times* visualization was created by John Burn-Murdoch and was based on data from March 23, 2020. To learn more about the decision made by the creators of the visualization, see their video.

4. Isaac Levy-Rubinett, "With Great Visualization Comes Great Responsibility," *Nightingale*, July 17, 2020, https://medium.com/nightingale/with-great-visualization-comes-great-responsibility-a863916d65c7. This explains how visualizations during the pandemic often acquired lives of their own.

5. A complete discussion of the use of data visualizations to mislead goes beyond the scope of this paper, but you can read more in Crystal Lee, Tanya Yang, Gabrielle Inchoco, Graham M. Jones,

and Arvind Satyanarayan, *Viral Visualizations: How Coronavirus Skeptics Use Orthodox Data Practices to Promote Unorthodox Science Online,* in *CHI Conference on Human Factors in Computing Systems: Making Waves, Combining Strengths,* online virtual conference originally planned for Yokohama, Japan, May 8-13, 2021.

6. I considered a link to such posts, but I did not want to dignify such nonsense with additional views.

7. The actual visualization provides an interactive scroll-over to display additional details about each county-level response. Josh Katz, Margot Sanger-Katz, and Kevin Quealy, "A Detailed Map of Who Is Wearing Masks in the US," *New York Times,* July 17, 2020, https://www.nytimes.com/interactive/2020/07/17/upshot/coronavirus-face-mask-map.html.

8. Edward R. Tufte, *The Visual Display of Quantitative Information* (Cheshire, CT: Graphics Press, 1983). Tufte introduces the idea of proportional ink more broadly, arguing, "The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities represented" (p. 56).