Health Matrix: The Journal of Law-Medicine

Volume 33 | Issue 1 Article 7

May 2023

A Thousand Views of the Cathedral: The Law, Politics, and Statistics of Pandemic Dashboards

Jeff Lingwall
Boise State University

TJ Bliss
Idaho State Board of Education

A THOUSAND VIEWS OF THE CATHEDRAL: THE LAW, POLITICS, AND STATISTICS OF PANDEMIC DASHBOARDS

Jeff Lingwall and TJ Bliss†

ABSTRACT

This Article explores the law, politics, and statistics of communicating data through the thousands of state, county, school district, and higher-education dashboards created in response to the COVID-19 global pandemic. Using a nationally distributed questionnaire and series of interviews with dashboard managers and stakeholders, we offer a wide-ranging view of data visualization practice in response to COVID-19. We pair this evidence with a survey of almost 3,000 entities responsible for public health communication, which resulted in collection of over 1,100 COVID-19 dashboards from a spectrum of government actors and private parties. We evaluate how legal issues were perceived and acted on, how data were politicized, and what technical challenges dashboard creators faced. We examine factors that led to creation of dashboards along with theory that explores the cognitive perception of dashboard elements. We explore the role of resources, and how those interplayed with specific software choices. We also examine the effect of local characteristics on dashboard creation, such as the presence of a county dashboard on data communication from local schools, the role of broadband access, technical occupations in the area, and

† Jeff Lingwall is an Assistant Professor of Legal Studies at Boise State University; PhD Carnegie Mellon University, JD Yale Law School. TJ Bliss is the Chief Academic Officer for the Idaho State Board of Education; PhD Brigham Young University. Both are founding partners at Redfish Metrics, LLC, through which the William and Flora Hewlett Foundation sponsored this research. Both authors aided in developing the Idaho state COVID-19 dashboard, which is included in the data analysis below. We thank R. John Azimi and faculty at the College of Business and Economics at Boise State University for helpful comments. We thank Kate Ferris and research assistants at Redfish Metrics, LLC for assistance. All views and any errors are our own.

so on. We conclude by suggesting a series of policies and practices that can be implemented to prepare for future data-based communication in a future public health crisis. These include practices to prepare for data visualization, enhanced resources devoted to public health communication, and the management of legal and political issues surrounding publicizing health information.

CONTENTS

> *"This article is meant to be only* one *of Monet's paintings of the Cathedral at Rouen. To understand the Cathedral one must see all of them."*
>
> ~ Calabresi & Melamed[1]

1. Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 Harv. L. Rev. 1089, 1090 (1972) ("As Professor Harry Wellington is fond of saying about many discussions of the law, this article is meant to be only *one* of Monet's paintings of the Cathedral at Rouen. To

> *"[H]undreds of government departments . . . had to quickly learn how to present data to the public as COVID-19 spread . . . . The resulting platforms — while sometimes flawed because of incomplete data, technology constraints or just awkward presentation — have become a mainstay of the pandemic."*
>
> ~ Marie Patino[2]

1. INTRODUCTION

The tragic spread of the SARS-CoV-2 virus[3] ("COVID-19") across the world created an unprecedented wave of attempts to publicly communicate data in a legally and politically fraught environment.[4] Due to the widespread accessibility of relatively inexpensive data visualization tools like Tableau,[5] PowerBI,[6]

understand the Cathedral one must see all of them.") (citing G. HAMILTON, CLAUDE MONET'S PAINTINGS OF ROUEN CATHEDRAL 4–5, 19–20, 27 (1960)).

2. *See generally* Carol M. Rose, *The Shadow of The Cathedral*, 106 YALE L. J. 2175, 2175 (1995) (describing the "venerable history" of the Calabresi and Melamed article).

3. Marie Patino, *The Rise of the Pandemic Dashboard*, BLOOMBERG CITYLAB (Sept. 25, 2021), https://www.bloomberg.com/news/features/2021-09-25/why-every-government-needs-a-covid-dashboard [https://perma.cc/69RA-AMRN].

4. In this manner, individuals in the United States experienced not one, but a multitude of competing, and sometimes conflicting, views of the state of the pandemic. A twist on the common legal-academic metaphor of "one view of the cathedral" seems apt to describe this simultaneous and multitudinous attempt to visualize an issue. *See supra* note 1.

5. *See, e.g.*, Inseok Ko & Hyejung Chang, *Interactive Visualization of Healthcare Data Using Tableau*, 23 HEALTHCARE INFORMATICS RES. 349 (2017) (describing the use of Tableau to visualize healthcare data).

6. *See* Steven M. Graves & Lauren He, *COVID-19 Mapping with Microsoft Power BI*, TERRA DIGITALIS (2020), http://terradigitalis.igg.unam.mx/html/ojs3/index.php/terra_digitalis/article/view/74 [https://perma.cc/622E-EMFP] (noting the use of PowerBI during the COVID-19 pandemic).

Smartsheets,[7] and ArcGIS,[8] entities in the United States such as state public health departments, higher education institutions, local education agencies ("school districts" or "districts"), and media organizations all began sharing pandemic-related data. These entities concurrently began navigating the laws surrounding public health data disclosure and the politics of the pandemic. This massive, simultaneous, and parallel development of data visualizations focused on a common crisis presents a unique opportunity to study mass communication of public health data and its regulatory environment.[9] Through studying why and how some entities chose to communicate data to the public, what legal challenges they faced, and what lessons they learned, organizations may be better prepared for future public health crises.[10]

Communicating data about COVID-19 has been fraught with legal, political, and technical concerns. Thousands of individuals, many having little experience with data visualization, found themselves creating, managing, or referring to data-driven

7. *See, e.g.*, DANIEL JACOB MELLER, A RESOURCE-EFFICIENT MODEL FOR PRESENTATION AND EVALUATION OF COVID-19 EPIDEMIOLOGICAL DATA FOR COUNTY PUBLIC HEALTH DEPARTMENTS (2020).

8. *See, e.g.*, Chris DuClos et al., *Peer Reviewed: Mapping Chronic Disease Risk Factors with ArcGIS Online in Support of COVID-19 Response in Florida*, 18 PREVENTING CHRONIC DISEASE (2021), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8091946/ [https://perma.cc/J4BT-CXLW].

9. The spread of COVID-19 across the United States occurred very quickly, and entities throughout the year began developing data communication tools. As described below, we collected dashboards during 2021 and 2022 itself or by using archived 2021 sources, and so the data do not reflect differences in timing of development during these years. In addition, as we examined some entities, they had already *closed* their COVID-19 dashboards at the time data were collected during 2021 and 2022. Some entities noted they had dashboards removed during the summer, when much of these data were collected. These entities were revisited after the summer, and if a dashboard was re-established this was counted in the data.

10. As of the date of this writing, COVID-19 and its Delta and ongoing Omicron variants continue to result in high caseloads, and so the results discussed below may be applicable both to the future *and* present pandemic responses.

dashboards to make decisions related to COVID-19.[11] Those creating the dashboards had to navigate potentially complicated legal requirements for what could, should, or had to be posted. They also faced a fraught political landscape in which individuals of widely varying political beliefs responded to public health requirements or recommendations in often extraordinarily passionate and negative ways.[12] Decision makers, from public health agencies to school districts, to the parents of children in those districts, attempted to incorporate a flood of information and quickly shifting policies from varying sources into their decision-making processes.[13]

To explore these complex dynamics, we asked public health officials, including dashboard creators and managers, across the United States about their experiences through an online questionnaire.[14] We then engaged in a series of in-depth interviews to examine survey responses in greater detail, drawing deeply on the expertise of officials directly engaged in communicating data about COVID-19. In order to better understand the scope and scale of COVID-19 communication, we also collected and

11. For a general description of the government perspective surrounding COVID-19 dashboards, *see* Patino, *supra* note 3.

12. On the other hand, the act of *removing* a dashboard could result in similar pushback! *See* Michael Lauzardo, *We Stopped Tracking Coronavirus Cases at the University of Florida. Here's Why.*, WASH. POST (Jan. 18, 2022), https://www.washingtonpost.com/outlook/2022/01/18/university-florida-covid-dashboard/ [https://perma.cc/YR7U-JSUU] ("On Dec. 31, we stopped updating the dashboard because I concluded that the numbers we were posting were no longer useful. This generated complaints, including accusations that we were covering up cases at a time when the situation was worsening by some measures."); *id.* ("Higher-education reporters took note of our move — as did our community, sometimes vociferously.").

13. One critique of the entire notion of a COVID-19 "dashboard" is the concept that dashboards, with their name taken from the instruments of a vehicle, are designed to be tools of *action*, yet most viewers of COVID-19 dashboards had little ability to "act" on the pandemic. Jacqueline Wernimont, *COVID-19 Dashboards are Vital, Yet Flawed, Sources of Public Information*, WASH. POST (Jan. 26, 2021), https://www.washingtonpost.com/outlook/2021/01/26/covid-dashboards-are-vital-yet-flawed-sources-public-information/ [https://perma.cc/PD82-TJBM].

14. We focus on the United States, while similar issues played out worldwide. *See, e.g.*, Patino, *supra* note 3.

analyzed hundreds of dashboards from every state and from samples of counties, colleges, and school districts across the United States.[15] In total, we examined almost 3,000 of these entities and found over 1,100 dashboards. We also examined those entities within our samples which did *not* produce dashboards, resulting in the first (of which we are aware) quantitative description of factors leading to COVID-19 data visualization by multiple entities across the United States.[16]

The picture that emerged from the questionnaire responses, interviews, and dashboard survey is complex. Respondents noted the challenge of constructing data visualization systems without expertise and while navigating strong partisan tensions.[17] Finding staff to review data, attempting to understand federal and state messaging, and creating visualizations which conveyed sufficient data in an understandable manner were all challenging tasks.[18] Even while we asked about political efforts to influence information on dashboards, with almost half of respondents noting their efforts had been "influenced" by political actors, respondents also reflected on politically-driven *responses* to dashboard data which challenged their efforts.[19] Despite this, respondents generally believed that visualizations did convey generally accurate and useful information about the state of the pandemic, although they were somewhat more skeptical about preparedness to communicate such data in the future.[20]

Based on our analysis of dashboards created—or not—across the country, the resources[21] available to entities appear to have

15. We include the District of Columbia when discussing states, but not other territories or other non-state jurisdictions.

16. In this Article, we offer a series of essential results and statistics from analyzing this database of dashboards. Future work may offer a more in-depth econometric examination of these data beyond the scope of traditional law review articles. We provide access to the dashboards collected below in *infra* note 131 in order to facilitate future research.

17. *See infra* note 120 and accompanying text.

18. *See infra* note 122 and accompanying text.

19. *See infra* note 121 and accompanying text.

20. *See infra* note 122 and accompanying text.

21. For the purposes of this Article, "resources" refers to money, staff, and technological expertise available to an entity. As described

played an important role in whether they created a dashboard.[22] Despite this trend, smaller, likely less well-resourced organizations often created complex, sophisticated representations of public health data, while many larger, more well-resourced organizations did not.[23] Regression analysis shows this relationship for school districts and higher education institutions. When comparing the creation of specific dashboard elements, there also appears to be a positive relationship between the presence of a county dashboard and whether school districts in those counties invested in a dashboard that went beyond simply reporting a table of known cases.[24] In contrast, for example, there is generally an insignificant relationship between the share of those particularly susceptible to COVID-19 within a county the creation of dashboards.[25] There is also a generally insignificant relationship between the creation of a dashboard and differences in political preferences within states.

Finally, the software used varied across entities. Counties, which often employ GIS software to track public works issues, often leveraged ArcGIS, while states with larger administrative resources often used Tableau, which is often considered a "premium" option for data visualization. In contrast, smaller entities such as school districts and higher education institutions favored basic HTML, in which it was more difficult to convey information through graphics or interactive elements. Because software capability is strongly related to the type of information

below, observable indications of resources are limited in this study to the size of entities and area economic characteristics.

22. *See infra* Table 2, Column (5) and accompanying text.

23. *See infra* Table 2, Column (6) and accompanying text.

24. We distinguish between simply reporting COVID-19 case counts in tabular form and the attempt to communicate statistical information to dashboard viewers through graphics or set-apart key indicators (e.g., total cases within the district, the percent change in total cases over time, and so on).

25. We examine the share of county population over age 85 as particularly susceptible to COVID-19. *See infra* Table 6 and discussion in the text. As discussed in Section 4, making strong causal claims from the regressions would be unwarranted because of the observational data collected. To help control for potential bias, we compare entities to other entities within the same state and control for a variety of variables potentially related to dashboard creation.

conveyed to the public,[26] it is significant that many entities with substantial resources—such as large universities—opted to communicate pandemic data without taking advantage of business intelligence services and their enhanced ability to create dashboard elements that would quickly and effectively communicate the state of the pandemic to their shareholders.[27]

Based on these data, we offer a range of concrete, practical proposals to prepare for data visualization needs in ongoing or future public health crises. First, legal frictions in the use of such data should be minimized in any entity's preparedness plan. Entities should remain up-to-date on the current state of public health reporting requirements, note what the Health Insurance Portability and Accountability Act (HIPAA) does and does not require, and create software systems that are flexible to capture changes in what legally can or cannot be reported, and from what source(s).[28] Second, having policies in place regarding specific visualizations themselves can be helpful, considering which data needs to be publicly reported, which key metrics are needed, and how information will be conveyed. Third, formulating a personnel process for data reporting and communication ahead of time—including who creates the key metrics to publish, the process by which those metrics will be reviewed, and so on—would help minimize the "we built the plane as we were flying it" critique of the data visualization process. Fourth, organizations need to create resources for data communication and foster cultures in which sharing data between entities in appropriate ways is permissible and done without being politicized. Finally, lowering the barriers to entry towards more sophisticated data visualizations may help those entities with fewer specialized resources. A significant example would be the creation of openly-licensed resources for non-experts to create graphics and interactive visualization elements.[29]

The next Section provides background on public health data communication in general, the creation of COVID-19 dashboards, and the legal environment surrounding public health communication. It also describes the development of business

26. *See infra* Table 3.

27. *See infra* note 33 and accompanying text.

28. *See infra* note 171 and accompanying text.

29. *See infra* note 172 and accompanying text.

intelligence software used to visualize pandemic data and examines three notable crises involving data visualization from the state-level COVID-19 dashboards developed by Florida, Georgia, and New York. The third Section then describes our questionnaire, interview process, and results. The fourth Section describes our quantitative study of dashboards in government, higher education, and K-12 education, and the results. The final Section concludes with discussion about the results of these studies and specific suggestions for preparing for future visualization of pandemic or other public safety crisis data.

2. BACKGROUND

This Section develops four background areas to provide context to the questionnaire, interview, and survey results presented later. First, *why* the need for such massive amounts of data collection and sharing; second, the legal environment of that communication; third, technical advancements in business intelligence services that made public communication through data visualization possible; and fourth, three examples of the interaction of law, politics, and data visualization through crises surrounding the Florida, Georgia, and New York COVID-19 dashboards.

A. The Demand for Public Communication of Health Data on COVID-19

The public demand for real-time, accurate data about COVID-19 spread throughout the United States almost as quickly as the virus itself. Civic leaders, education officials, and the public in general wanted information about quantitative aspects of the pandemic such as infection rates, hospitalizations, and death rates at nearly every level of jurisdiction, from the national to the very local.[30]

30. As the pandemic progressed, the ability to provide accurate numbers was increasingly challenged. *See* Lauzardo*, supra* note 12 ("[S]o many students and staff were forgoing the school's official testing site, and other sites that reported to us, in favor of at-home tests — and not reporting the results to us — that we lost confidence that our totals bore any relation to reality."); *id.* ("We also knew that many students were experiencing mild covid-19 symptoms but attributing them to allergies or a cold, and so not reporting them. Still others didn't test because they didn't want to

In July 2020, the United States Centers for Disease Control and Prevention (CDC) noted this demand and began providing specific support to state-level health departments in their efforts to develop state-level COVID-19 dashboards. Debra Lubar, deputy director of management and operations for the National Center for Emerging and Zoonotic Infectious Diseases at CDC noted:

> [T]here's a real demand for dynamic visualizations of all kinds of data in all of our emergency responses, and particularly in this one that's affecting every US community but in a way that differs across the country. There are critical data elements that are really important for people to see about their communities, where their families live, or places they plan to travel.[31]

Demand for easy-to-digest visual data about COVID-19 likely grew from the desire for safety in daily-decision making. Theoretically, a quick look at a local dashboard could help someone decide whether to send a child to school, determine the best time to travel to another state to visit loved ones, or determine when it was safe to allow employees to return to the office.[32]

---

be barred from dining rooms and classrooms — a trend that increased over the course of the pandemic.").

31. *See* Este Geraghty, *COVID-19: CDC Supports State Dashboards to Better Monitor Cases and Capacity.* ESRI BLOG (July 9, 2020), https://www.esri.com/about/newsroom/blog/covid-19-cdc-supports-state-dashboards-to-better-monitor-cases-and-capacity/ [https://perma.cc/K3RM-BPR2].

32. One year into the pandemic Jacqueline Wernimont, Distinguished Chair of Digital Humanities and Social Engagement at Dartmouth College, wrote about the individual-level utility of COVID-19 dashboards:

> [T]he dashboard is one common way for millions of people to engage with and make decisions about the pandemic . . . [but] they fail to function in some key ways. In fact, covid-19 dashboards are remarkably ill-suited to the complexities of our pandemic reality . . . because historically the dashboard has been designed for a reader or viewer who has some power to act . . . Today's dashboard visualizations tend to operate at much larger scales, and are far more complex and frequently updated. But in some ways they are perhaps less useful to ordinary people . . . For example, there are data lags of up to 10

Civic and education leaders, in particular, relied on COVID-19 dashboards to meet this demand in their local communities. Counties, cities, higher education institutions, and school districts developed data visualizations as a way to communicate local health data in a universally accessible and dynamically accurate manner[33], with new dashboard development continuing as the pandemic continued.[34]

Developing visualizations that were accurate and accessible was a challenging task. Because most dashboard consumers are not data scientists, they may not find it rational to devote substantial time parsing a large amount of information before making a personal health decision—thus how data are aggregated, analyzed, and presented matters crucially to their capacity to affect behavior. A large body of research supports

> days in digital dashboards that are difficult for users to wrap their heads around. Additionally, many dashboards provide global or national level data but not the local data that could actually inform an individual decision about where to go. Some local- and county-level data are available, but not always in the user-friendly format of the dashboard, and certainly not consistently across the country . . . Indeed, public health dashboards, like our many covid-19 dashboards, are unusual in the history of dashboards in that they share information but not in a way that allows ordinary people to take action. We might well wish that covid-19 dashboards functioned as both protective barriers and a way to see the information we need to steer ourselves clear of trouble. However, there is little about our current situation that suggests that any one of us are in the driver's seat of a global health vehicle.

Wernimont, *supra* note 13.

33. *E.g.*, Jessica Kent, *Big Data Dashboards to Help US Colleges Track COVID-19 Spread*, HEALTH IT ANALYTICS (Aug. 9, 2020), https://healthitanalytics.com/news/big-data-dashboards-to-help-us-colleges-track-covid-19-spread [https://perma.cc/3KXF-R2AE] (remarking on Boston University's dashboard and noting that "[l]eaders will update the dashboard daily, and will include daily and cumulative numbers as well as seven-day averages and comparative figures from the local county. The university expects that the dashboard will enable proactive measures to stop further spread of the virus.").

34. *See Greenville County Schools Creates COVID-19 Dashboard,* WYFF NBC NEWS (Aug. 19, 2021), https://www.wyff4.com/article/greenville-county-schools-creates-covid-19-dashboard/37347716 [https://perma.cc/G4DW-JRLX].

this. For example, consider the concept of "cognitive fit." Cognitive fit occurs when "problem presentation" and "problem solving tools" match "the characteristics of the task."[35] For instance, researchers have often considered whether graphs or tables are better able to convey quantitative information.[36] In part, this depends on the particular "information task" at hand. If the information to be gleaned is about relationships in the data, graphics may be more helpful than tables. When asking whether County X has lower disease rates than Counties Y and Z, a visual comparison of disease rates could convey this information quickly and memorably, showing which county has lower rates, and by how much.[37] In contrast, if the information task involves extracting distinct or discrete data of interest, such as answering what the disease rate at County X *is*, this is facilitated by tabular information.[38]

There are many ways differences in data can be visualized, and so the choice of how to convey quantitative information is complex. Assuming the information task is comparison-based, such as in the first example above, and that graphical representation of quantitative data is a better cognitive fit than tabular representation, what *kind* of graph would be best has been extensively studied. For example, similar to theory on cognitive fit, interpreting graphical data can be thought of as a taxonomy of skills, beginning with elementary skills required to "extract information from the data," intermediate skills that seek to "find relationships in the data," and overall skills that "move beyond the data," such as predictive inference related to the graph.[39] The ability of consumers of information to accomplish these tasks

35. Iris Vessey & Dennis Galletta, *Cognitive Fit: An Empirical Study of Information Acquisition*, 2 INFO. SYS. RES. 63, 63 (1991).

36. This type of analysis has been conducted for decades. *E.g.*, John Noble Washburne, *An Experimental Study of Various Graphic, Tabular, and Textual Methods of Presenting Quantitative Material*, 18 J. EDUC. PSYCH. 361, 361 (1927) (studying how school children reacted to various presentations of identical quantitative information).

37. *Id.* at 68.

38. *Id.*

39. Susan N. Friel, Frances R. Curcio & George W. Bright, *Making Sense of Graphs: Critical Factors Influencing Comprehension and Instructional Implications*, 32 J. MATH. EDUC. 124, 130–31 (2001).

depends on the choice of graphical form. For example, while they might convey the same information, bar graphs are more easily interpreted than pie charts, as the judgment involved in noting height is easier than in noting angle.[40]

For example, in one taxonomy of how accurately consumers interpret graphical data, "Position on a common aligned scale" (such as a line graph) conveys information most likely to be accurately perceived, followed by "Position on common unaligned scales," (such as a scatter plot), "Length" (such as segmented bar charts), "Angle/slope" (as in pie charts), "Area" (as in circle or blob charts), "Volume" (as in three-dimensional charts), and "Color hue" (as when maps are color coded).[41] At times, great depth of information can be conveyed by combining these aspects of data representation. Scatter plots in which the size or area of different-sized points conveys information may be effective in showing a third aspect of the data not reflected by the x- and y-axes themselves. Similarly, color may differentiate groups in line plots or scatter plots, and so on.

Another challenging aspect in COVID-19 data visualization is the possibility of combining spatial and temporal elements. The multidimensional aspect of these data can be difficult to represent in static form. This challenge has existed since the beginning of graphical representation of data. When Galileo began his observations of Jupiter and its moons, he published his findings in the noteworthy *The Starry Messenger*.[42] He had observed "small, bright starlets" (Jupiter's moons) which moved about the planet at different times of observation.[43] He mapped the spatial arrangement of Jupiter and the moons over time in a series of charts depicting the location of the bodies with dots and circles over a series of nights, communicating in graphical form what he observed.[44] We take for granted the ease with which we create and interpret much statistical information today, but it took 300

40. *Id.* at 135.

41. *Id.* at 136 (presenting Cleveland and McGill's "Taxonomy of Specifiers Ordered from Most to Least Accurately Used", citing W.S. CLEVELAND, THE ELEMENTS OF GRAPHING DATA (1985); C.M. Carswell, *Choosing Specifiers: An Evaluation of the Basic Tasks Model of Graphical Perception*, 34 HUM. FACTORS 535 (1992)).

42. *See* TUFTE, ENVISIONING INFORMATION 97–101 (1990).

43. *Id.* at 97.

44. *Id.* at 97–98.

years for Galileo's charts to be updated with smooth curves which showed the location of the satellites at each point in time.[45] In many ways, the graphical representation of data combines statistical literacy with aspects of graphic design, none of which may be familiar to non-experts tasked with creating COVID-19 visualizations.

B. The Legal Environment of Public Health Data Communication

In addition to navigating new complexities in visualizing data to their stakeholders, dashboard creators face a range of unfamiliar *legal* challenges when reporting public health-related data to the public. These include navigating HIPAA, ever changing regulations governing data collection from hospitals, accessibility standards and regulations[46], and standards for what data should be communicated to the public.

HIPAA's Privacy Rule is well familiar—if not well understood—by patients who have received medical services in the United States.[47] The Privacy Rule is a regulation promulgated by the Department of Health and Human Services (HHS), under authority of HIPAA, to create standards for protecting health information in the United States.[48] For covered entities (health

45. *Id.* at 100.

46. *See* ADA NAT'L NETWORK, *An Overview of the Americans with Disability Act* (2017), https://adata.org/factsheet/ADA-overview [https://perma.cc/3LW3-HNP9]; W3C WEB ACCESSIBILITY INITIATIVE, *W3C Accessibility Standards Overview* (Feb. 2, 2022), https://www.w3.org/WAI/standards-guidelines/ [https://perma.cc/L8CR-YQJK].

47. *See* Health Insurance Portability and Accountability Act of 1996 (HIPAA), 42 U.S.C. § 1320d (1996). The signing of a boilerplate and likely unread HIPAA forms by medical patients is now routine practice in the United States. *Cf.* Mark Hochhauser, *Compliance vs. Communication: Readability of HIPAA Notices (Hochhauser)*, PRIVACY RIGHTS CLEARINGHOUSE (Nov. 1, 2003), https://privacyrights.org/resources/compliance-vs-communication-readability-hipaa-notices-hochhauser [https://perma.cc/QD5P-BXRK] (discussing whether HIPAA notices are written in "plain language" as required by statute and examining the implications of "unreadable" notices).

48. *See* U.S. DEP'T OF HEALTH AND HUM. SERVS. [hereinafter HHS], *Summary of the HIPAA Privacy Rule*, HHS.GOV (July 26, 2013), hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html [https://perma.cc/3GW4-JTHP] ("Sections 261 through 264 of HIPAA require the Secretary of HHS to publicize standards for

plans, care providers, clearinghouses, and their business associates), individual identifiable personal health information (PHI) cannot be used or disclosed except in accordance with the Rule.[49] In essence, the regulation then restricts the use of PHI generally and permits its use in a variety of limited, specific ways, such as for treatment and payment for health care operations, public interest activities (such as reporting data on domestic violence or abuse, or as required by a court order), and for purposes of public health activities or research (once specific identifiers have been removed).[50]

During the pandemic, many HIPAA-related issues became prominent, and two particular questions arose relating to communicating information to the public about the state of the pandemic.[51] First, could entities such as school districts, higher

---

the electronic exchange, privacy and security of health information."). HIPAA required the Secretary of HHS to create privacy rules if Congress did not do so within three years of passage. Congress did not act, and after an active notice and comment period with 11,000 comments received, the Privacy Rule became law on December 28, 2000. *Id.*

49. A health care clearinghouse is an entity that processes information between health care entities. *See* CTRS. FOR MEDICARE & MEDICAID SERV.'S [hereinafter CMS], *Are You a Covered Entity*, CMS.GOV (Dec. 1, 2021), https://www.cms.gov/Regulations-and-Guidance/Administrative-Simplification/HIPAA-ACA/AreYouaCoveredEntity [https://perma.cc/CKP5-PWU4] (defining covered entities under HIPAA).

50. HHS, *Summary of the HIPAA Privacy Rule*, *supra* note 48 ("A covered entity is permitted, but not required, to use and disclose protected health information, without an individual's authorization, for the following purposes or situations: (1) To the Individual (unless required for access or accounting of disclosures); (2) Treatment, Payment, and Health Care Operations; (3) Opportunity to Agree or Object; (4) Incident to an otherwise permitted use and disclosure; (5) Public Interest and Benefit Activities; and (6) Limited Data Set for the purposes of research, public health or health care operations.") (footnote omitted).

51. *See* Lydia Wheeler, *Confusion Over Health Privacy Law Seen Impeding Covid Battle*, BLOOMBERG LAW (Aug. 17, 2021), https://news.bloomberglaw.com/health-law-and-business/confusion-over-health-privacy-law-seen-impeding-covid-battle [https://perma.cc/TJQ3-K9YF] ("The 1996 law went viral on social media after politicians and pro-athletes . . . inaccurately cited it as they declined to answer questions about whether they've been vaccinated against COVID-19.").

education institutions, and employers collect individual or aggregated data on COVID-19 diagnoses, exposures, or vaccinations? Second, to what extent could those entities share that data within and outside of these organizations? The answer to the first question is generally yes—entities may freely collect information on COVID-19 issues from their stakeholders. In fact, the CDC and Equal Employment Opportunity Commission (EEOC) *encouraged* employers to inquire about COVID-19 related issues.[52] Contrary to some public misunderstanding, HIPAA limits the disclosure (by covered entities) of health information once collected, not the collection *of* health information from individuals *per se*.[53] While the Americans with Disabilities Act (ADA)[54], not HIPAA, would require such information be kept confidential by employers, neither the ADA nor HIPAA prohibit its collection.[55]

The answer to the second question—can entities share that data once collected—is also generally yes, particularly for aggregated data. First, many of the entities collecting data are not healthcare providers, plans, or clearinghouses, and hence do not qualify as "covered entities" to which HIPAA applies.[56]

52. *What Employers Need to Know About HIPAA*, NAT'L L. REV., (Mar. 23, 2020), https://www.natlawreview.com/article/what-employers-need-to-know-about-hipaa [https://perma.cc/AN9U-SR5M].

53. In addition, HIPAA only limits disclosure to "covered entities", which excludes the vast majority of entities with interest in gathering information on COVID-19 status, such as employers. One wrinkle is when a private employer also operated as a self-insured health plan. In that case, if the information was gathered pursuant to the plan, the information would be controlled by HIPAA. *Id.*

54. *See supra* note 46.

55. Many follow-up questions might be unlawful, such as inquiring *why* a particular individual had chosen not to be vaccinated. *See* HIPAA JOURNAL, *Is It a HIPAA Violation to Ask for Proof of Vaccine Status?*, HIPAA J. (May 25, 2021), https://www.hipaajournal.com/is-it-a-hipaa-violation-to-ask-for-proof-of-vaccine-status/ [https://perma.cc/PAW2-466E].

56. *Id.* ("HIPAA only applies to HIPAA-covered entities — healthcare providers, health plans, and healthcare clearinghouses — and their business associates. If an employer asks an employee to provide proof that they have been vaccinated in order to allow that individual to work without wearing a facemask, that is not a HIPAA violation as HIPAA does not apply to most employers.").

Second, even when HIPAA does apply, HIPAA specifically allows sharing of PHI to "public health authorities" for "public health activities" such as aggregating data for reporting infection rates.[57] HHS Office for Civil Rights issued a memorandum in March 2020, which outlined this principle, noting that (1) "HIPAA permits a covered entity to disclose PHI to a public health authority (such as the Centers for Disease Control and Prevention (CDC), or state, tribal, local, and territorial public health departments) that is authorized by law to collect or receive PHI for the purpose of preventing or controlling disease . . . ."[58] Third, in March 2020, HHS also issued a limited waiver of HIPAA sanctions and penalties due to the pandemic, such as the penalties for not honoring a request to opt out of the directory of a facility.[59] While "[s]omehow over the years, HIPAA has wrongly morphed into a generic term for health privacy," together, these principles and regulations limited many broad concerns about potential HIPAA violations resulting from creating and publishing dashboards that reported aggregated, non-personally-identifiable information about COVID-19.[60]

57. *See* HHS, *Summary of the HIPAA Privacy Rule*, *supra* note 48. Public health authorities are an authority of the United States government, a state, a political subdivision of a state, and so forth. 45 C.F.R. §§ 164.501–164.512.

58. HHS: OFF. C.R., *COVID-19 and HIPAA: Disclosures to Law Enforcement, Paramedics, Other First Responders and Public Health Authorities*, https://www.hhs.gov/sites/default/files/covid-19-hipaa-and-first-responders-508.pdf [https://perma.cc/8SYT-4YKS] (last visited Feb. 8, 2022) (released at HHS, *OCR Issues Guidance to Help Ensure First Responders and Others Receive Protected Health Information about Individuals Exposed to COVID-19*, HHS.GOV (Mar. 24, 2020), https://www.hhs.gov/about/news/2020/03/24/ocr-issues-guidance-to-help-ensure-first-responders-and-others-receive-protected-health-information-about-individuals-exposed-to-covid-19.html [https://perma.cc/FDJ8-WYRP]) (listing "public health surveillance," "public health investigations," and "public health interventions" as permissible purposes for which PHI may be released without permission).

59. HHS, *COVID-19 & HIPAA Bulletin: Limited Waiver of HIPAA Sanctions and Penalties During a Nationwide Public Health Emergency* (Mar. 2020), https://www.hhs.gov/sites/default/files/hipaa-and-covid-19-limited-hipaa-waiver-bulletin-508.pdf [https://perma.cc/ZCU2-7J8M].

60. Wheeler, *supra* note 51.

A related legal issue was how hospitals and nursing homes reported COVID-19-related data. Throughout 2020 and 2021, hospitals faced a "maddening flurry" of changing requirements in reporting, such as how often data should be reported, to whom reports would go, and what data was collected.[61] In March 2020, Vice President Mike Pence asked for daily reports from hospitals, through the National Healthcare Safety Network to the CDC, including information on COVID-19 testing and hospital capacity.[62] While not a requirement, about 85% of hospitals reported data.[63] Then, a formal legal framework for how these entities report data to the federal government was created through the Coronavirus Aid, Relief, and Economic Security Act (CARES Act), and regulations promulgated by the Centers for Medicare & Medicaid Services (CMS) soon followed.[64]

61. Pien Huang & Selena Simmons-Duffin, *Trump Administration Plans Crackdown on Hospitals Failing to Report COVID-19 Data*, NAT'L PUB. RADIO (Sept. 24, 2020), https://www.npr.org/sections/health-shots/2020/09/24/916310786/trump-administration-plans-crackdown-on-hospitals-failing-to-report-covid-19-dat [https://perma.cc/WLY8-YMRZ] ("The reporting system drew national attention in July when the Department of Health and Human Services told hospitals to stop reporting information — such as the number of COVID-19 patients and the availability of intensive care beds — to the Centers for Disease Control and Prevention, and instead report it into a new system managed directly by HHS, the CDC's parent agency. The switch raised concerns from politicians and public health experts about the sidelining of the CDC, the nation's public health agency, in the midst of a pandemic.").

62. Pien Huang & Selena Simmons-Duffin, *COVID-19 Guidance for Hospital Reporting and FAQs*, NAT'L PUB. RADIO (July 31, 2020, 5:00 AM), https://www.npr.org/sections/health-shots/2020/07/31/897429054/covid-19-hospital-data-system-that-bypasses-cdc-plagued-by-delays-inaccuracies [https://perma.cc/N924-V45W] ("[T]he public data hub created under the new system is updated erratically and is rife with inconsistencies and errors, data analysts say.").

63. Huang & Simmons-Duffin, *Trump Administration Plans Crackdown on Hospitals Failing to Report COVID-19 Data*, *supra* note 61.

64. *See* CMS, *Trump Administration Strengthens COVID-19 Surveillance with New Reporting and Testing Requirements for Nursing Homes, Other Providers*, CMS.GOV (Aug. 25, 2020), https://www.cms.gov/newsroom/press-releases/trump-

In April 2020, CMS asked governors to create plans for COVID-19 testing in nursing homes, and in May 2020, CMS provided recommendations for baseline testing in nursing homes.[65] CDC admitted that "in tracking the spread of the virus, it had been combining tests that detect active infection with those that detect recovery from COVID-19. That system muddied the picture of the pandemic."[66] In June 2020, HHS required laboratories to report all COVID-related information to the Secretary of HHS.[67] Then in July 2020, HHS implemented a change to the rule that required direct reporting to HHS alone, rather than to HHS and the CDC, through an online portal run by an entity called TeleTracking.[68] This required hospitals to adapt to a new "data system" and stopped updates of data that CDC was publishing for researchers.[69] It also eliminated the CDC's role in creating estimates for the availability of Intensive Care Unit beds and vetting the reported data for accuracy.[70]

---

administration-strengthens-covid-19-surveillance-new-reporting-and-testing-requirements [https://perma.cc/TK9E-ULJK].

65. *Id.*

66. Sheryl Gay Stolberg, *Trump Administration Strips C.D.C of Control of Coronavirus Data*, N.Y. TIMES (July 14, 2020), https://www.nytimes.com/2020/07/14/us/politics/trump-cdc-coronavirus.html [https://perma.cc/844H-63PY].

67. Medicare and Medicaid Programs, Clinical Laboratory Improvement Amendments (CLIA), and Patient Protection and Affordable Care Act; Additional Policy and Regulatory Revisions in Response to the COVID-19 Public Health Emergency, 85 Fed. Reg. 54820 (Sept. 2, 2020).

68. HHS, *COVID-19 Guidance for Hospital Reporting and FAQs For Hospitals, Hospital Laboratory, and Acute Care Facility Data Reporting*, https://www.hhs.gov/sites/default/files/covid-19-faqs-hospitals-hospital-laboratory-acute-care-facility-data-reporting.pdf [https://perma.cc/34GF-XCL7] (last updated Jan. 2, 2023); Nicholas Florko & Eric Boodman, *How HHS's New Hospital Data Reporting System Will Actually Affect the U.S. COVID-19 Response*, STATNEWS (July 16, 2020), https://www.statnews.com/2020/07/16/hospital-data-reporting-covid-19/ [https://perma.cc/GV9P-2BKQ].

69. *Id.*

70. Pien Huang & Selena Simmons-Duffin, *COVID-19 Hospital Data System That Bypasses CDC Plagued By Delays, Inaccuracies*, NAT'L PUB. RADIO (July 31, 2020 5:00 AM), https://www.npr.org/sections/health-shots/2020/07/31/897429054/covid-19-hospital-

This was controversial. Details of how the contractor that created TeleTracking was chosen were unclear,[71] and these changes were implemented when the Trump White House was arguing with the director of the CDC over the response to the pandemic.[72] Further, the switch to TeleTracking was announced with little notice, and non-compliant hospitals could lose access to remdesivir, an antiviral medication used to treat COVID-19 patients.[73] The change took effect when hospitals were overwhelmed, and the HHS system experienced technical problems that caused errors and delays. This change in reporting was also costly to hospitals, which were already under cost-constraints related to loss of elective medical procedures during the pandemic. Moreover, hospitals were not compensated by the federal government for the altered reporting requirements.[74]

In August 2020, CMS amended and revised a number of regulations[75], including the requirement that hospitals report daily data such as the number of COVID-19 positive patients and

---

data-system-that-bypasses-cdc-plagued-by-delays-inaccuracies [https://perma.cc/V8DE-VN94].

71. *Id.* The bid to create the system was apparently noncompetitive. *Id.*

72. Pien Huang & Selena Simmons-Duffin, *White House Strips CDC of Data Collection Role for COVID-19 Hospitalization*, NAT'L PUB. RADIO (July 15, 2020 1:31 PM), https://www.npr.org/sections/health-shots/2020/07/15/891351706/white-house-strips-cdc-of-data-collection-role-for-covid-19-hospitalizations [https://perma.cc/HY6G-CTXA].

73. *See, e.g.*, JH Beigel *et al.*, *Remdesivir for the Treatment of Covid-19 — Preliminary Report*, 383 N. ENG. J. MED. 1813 (2020); Huang & Simmons-Duffin, *COVID-19 Hospital Data System That Bypasses CDC Plagued By Delays, Inaccuracies*, *supra* note 70.

74. *Id.* ("The July change to data reporting created a large and costly administrative burden for hospitals without providing funding to help them fulfill it. The new draft guidance would further expand the scope.").

75. *See* 42 C.F.R. §§ 410, 413, 414, 422, 423, 482, 483, 485, 488 & 493. These changes are outlined in a CMS regulatory document available at *Medicare and Medicaid Programs, Clinical Laboratory Improvement Amendments (CLIA), and Patient Protection and Affordable Care Act; Additional Policy and Regulatory Revisions in Response to the COVID-19 Public Health Emergency* (Aug. 24, 2020), https://www.cms.gov/files/document/covid-ifc-3-8-25-20.pdf [https://perma.cc/ZAD7-JWEX].

the levels of medical equipment such as ventilators,.[76] Laboratories testing for COVID-19 were similarly required to report individual patient results.[77] During September 2020, CMS mandated that hospitals report data seven days a week, including weekends, and threatened to remove Medicaid/Medicare funding from non-compliant hospitals.[78] Nursing homes were required to routinely test staff, and penalties would be assessed for failing to report weekly COVID-related data.[79] During October 2020, HHS announced that it would begin requiring hospitals to report flu data together with data on COVID-19, and as remdesivir supplies increased, hospitals were no longer required to report their holdings of the drug.[80]

Hospitals and other health-care entities were not the only entities with reporting requirements. Schools across the country faced patchwork state-by-state requirements for what data was suggested or required to be reported, and to whom.[81] As examples, in New York and Kentucky, schools were required to report data to the state's health department, and the state then published

76. Non-compliance with reporting requirements risked termination of Medicare and Medicaid payments. CMS, *Trump Administration Strengthens COVID-19 Surveillance with New Reporting and Testing Requirements for Nursing Homes, Other Providers*, *supra* note 64.

77. Non-compliance for laboratories risked a $1,000 a day fine. *Id.* Previous regulation required that laboratories report test results and accompanying data. *Id.*

78. CMS, *Trump Administration Strengthens COVID-19 Surveillance with New Reporting and Testing Requirements for Nursing Homes, Other Providers*, *supra* note 64.

79. 85 Fed. Reg. 54820, *supra* note 67.

80. *Interim Final Rule (IFC), CMS-3401-IFC; Requirements and Enforcement Process for Reporting of COVID-19 Data Elements for Hospitals and Critical Access Hospitals*, CTRS. FOR MEDICARE AND MEDICAID SERVS. (Oct. 6, 2020), https://www.cms.gov/files/document/interim-final-rule-ifc-cms-3401-ifc-requirements-and-enforcement-process-reporting-covid-19-data.pdf [https://perma.cc/2H99-8C86].

81. There was no federal requirement to monitor COVID-19 cases in schools. *See* Benjy Sarlin & Suzy Khimm, *Coronavirus is Spreading in Schools, but the Federal Government Isn't Keeping Count*, NBC NEWS (Aug. 17, 2020), https://www.nbcnews.com/news/us-news/coronavirus-spreading-schools-no-one-tracking-all-outbreaks-n1236964 [https://perma.cc/JM5C-4ETS].

the data on a statewide K-12 dashboard.[82] In Texas, schools were required to report COVID-19 data through an online form to the state's health department, as well as to their county health departments.[83] It was suggested, but not required, that schools share this information with parents.[84] Michigan imposed a requirement that "schools . . . disclose cases prominently on the school's website,"[85] while in Florida and California local reporting remained optional.[86] The next subsection then considers how

82. *Governor Cuomo Announces School Reporting Requirements and Online Dashboard to Track COVID-19*, GOVERNOR.NY.GOV (Sept. 3, 2020), https://www.governor.ny.gov/news/governor-cuomo-announces-school-reporting-requirements-and-online-dashboard-track-covid-19 [https://perma.cc/8ADR-R9FL]; Rhonda Miller, *New K-12 Covid -19 Dashboard Live for Kentucky Schools to Self-Report Data*, WKMS (Oct. 5, 2020), https://www.wkms.org/government-politics/2020-10-05/new-k-12-covid-19-dashboard-live-for-kentucky-schools-to-self-report-data [https://perma.cc/82A7-S63L].

83. Brian Lopez, *Texas Says Schools Now Must Notify Families of Positive COVID-19 Cases in Classrooms*, TEXAS TRIBUNE (Aug. 19, 2021), https://www.texastribune.org/2021/08/19/texas-schools-covid-19-cases/ [https://perma.cc/A7Q8-LJ53] ("Districts must still report positive cases to their local health departments and the state. Local health officials are also allowed to investigate COVID-19 cases in schools.").

84. As one local article stated, "Experts have emphasized that releasing public data is critical for families — as well as political leaders — to assess the risks and rewards of sending students back to school." Alisa Swaby & Emma Platoff, *State Releases Numbers Showing Low Texas Public School Infection Rates, But the Data is Limited*, TEXAS TRIBUNE (Sept. 17, 2020), https://www.texastribune.org/2020/09/17/texas-students-coronavirus-schools/ [https://perma.cc/E6N5-TCNW].

85. Ron French, *Michigan Re-Orders Schools to Disclose COVID Cases After Supreme Court Rulings*, BRIDGE MI (Oct. 6, 2020), https://www.bridgemi.com/talent-education/michigan-re-orders-schools-disclose-covid-cases-after-supreme-court-ruling [https://perma.cc/38CH-X793].

86. *See* Travis Gibson & Joe McLean, *Still No Word on When State plans to Publish COVID-19 Data for Schools*, NEWS4JAX (Sept. 12, 2020), https://www.news4jax.com/news/local/2020/09/11/still-no-word-on-when-state-plans-to-publish-covid-19-data-for-schools/ [https://perma.cc/3ZD9-WHJ5]; Kristen Taketa, *San Diego County Had 283 School COVID-19 Cases in Two Months, But Data Are Limited*, SAN DIEGO UNION TRIBUNE (Oct. 9, 2020), https://www.sandiegouniontribune.com/news/education/story/20

many of these entities were able to disclose COVID-19 information to the public.

C. The Development of BI and GIS Programs

Organizations that desired or were required to communicate pandemic-related data were able to take advantage of recent advancements in visual data communication targeted at users without extensive technical backgrounds. The burgeoning field of user-friendly business intelligence tools was well-suited to be adapted by school districts, higher education institutions, and local governments who possessed comparatively little institutional expertise in data analytics. Consider the development of three such tools: Tableau, PowerBI, and ArcGIS. Tableau was created by a trio of Stanford University researchers in 2003, as an outgrowth of a PhD dissertation on "automatic design of graphical presentations."[87] They developed a suite of tools that allowed users to engage with data from a wide variety of sources through a predominantly visual, drag-and-drop interface, rather than writing extensive lines of computer code. As users engaged with data, they were able to quickly create interactive, easily-updated, online-accessible visualizations without the need to engage computer scientists or other information technology professionals.[88] Tableau grew at a tremendous rate, reaching hundreds of millions of dollars in sales by 2020.[89]

20-10-09/san-diego-county-had-283-school-covid-19-cases-in-two-months-but-data-are-limited [https://perma.cc/77VU-W37R].

87. John Mackinlay, *Analyzing the History of Tableau Innovation*, TABLEAU (Dec. 2, 2021), https://www.tableau.com/about/blog/2021/12/analyzing-history-tableau-innovation [https://perma.cc/P75Y-6VZS].

88. "Interactive" elements include, for example, the ability to hover over elements of a dashboard to view definitions or further data. *See COVID-19: Activity Level by Region and County*, WISC. DEP'T HEALTH SERVS. (Jan. 20, 2022), https://www.dhs.wisconsin.gov/covid-19/local.htm [https://perma.cc/6BQ6-VZ8A] (describing how to interact with Wisconsin's state COVID-19 dashboard and pages).

89. *See, e.g.*, *Tableau Reports First Quarter 2019 Financial Results*, TABLEAU.COM (May 2, 2019), https://www.tableau.com/about/press-releases/2019/tableau-reports-first-quarter-2019-financial-results [https://perma.cc/JYH9-NZ44].

Perhaps in response to the business intelligence skillset allowed by Tableau, Microsoft responded by developing its own suite of dashboard creation tools for business intelligence, called PowerBI.[90] This suite, while perhaps not allowing the extended customizability and visual appeal of Tableau, integrated with Microsoft's widespread suite of office tools, including Excel.[91] It was first named Project Crescent and made available to the public in 2011, rebranded as PowerBI for Office 365 in 2013, and made available to the general public in 2015.[92] Many of the visualizations in PowerBI began life as "Datazen" (a Canadian company Microsoft had acquired) visualizations.[93]

Finally, contrast these business-intelligence-focused software packages with ArcGIS, a powerful geographical information system (GIS) tool widely used by local governments, and drawn on extensively during the pandemic to represent data. The concept of GIS is a database and set of tools which specialize in cartographic data.[94] These began in the 1960s as computing power began to enter the academic world and governments wished to track resources and manage projects.[95] The company ESRI (Environmental Systems Research Institute) was founded in the late 1960s, and developed a commercial arm in the 1980s which produced ArcGIS, a popular GIS platform.[96] Because many government entities already had extensive experience with GIS products, using those tools to visualize COVID-19 data was a

90. *See* David Taylor, *Power BI Tutorial: What is Power BI? Why Use? DAX Examples*, GURU99 (Jan. 15, 2022), https://www.guru99.com/power-bi-tutorial.html [https://perma.cc/4TGG-AU5F].

91. *Id.*

92. *Id.*

93. Mark W. Kaelin, *Microsoft Adds Datazen and Mobility Apps to Power BI*, TECH REPUBLIC (May 19, 2015), https://www.techrepublic.com/article/microsoft-adds-datazen-and-mobility-apps-to-power-bi/ [https://perma.cc/V4GT-EURZ].

94. *See* KANG-TSUNG CHANG, INTRODUCTION TO GEOGRAPHIC INFORMATION (2018).

95. *History of GIS*, ESRI, https://www.esri.com/en-us/what-is-gis/history-of-gis [https://perma.cc/6M5F-F585] (last visited Feb. 8, 2022).

96. *Id.*

natural fit, particularly as ESRI provided COVID-specific tutorials on dashboard creation.[97]

D. Crises in COVID-19 Data Visualization

The potential politicization of issues related to the pandemic has been the subject of study from a wide variety of perspectives, such as news coverage,[98] health behaviors,[99] racism,[100] state executive orders,[101] and so on.[102] In particular, data collection and presentation processes themselves became subject to significant political-based controversy. Examples from Florida, Georgia, and New York are illustrative. The cause-and-effect relationship between political-based actions and resulting controversies is difficult to establish, but these examples illustrate the significant potential for political-based motives, rather than public-health rationales, to influence public representation of pandemic-related data.

In Florida, which COVID-19 data would be reported, and by whom, became the subject of popular press, criminal prosecution, and news headlines. State GIS specialist Rebekah Jones claimed that she was fired when state officials requested she "manually

97. Vicki Lynn Cove, *Visualize Global COVID-19 Trends in ArcGIS Insights*, ARCGIS (Aug. 4, 2021), https://learn.arcgis.com/en/projects/visualize-global-covid-19-trends-in-arcgis-insights/ [https://perma.cc/8TRN-N4RY].

98. P. Sol. Hart, Sedona Chinn & Stuart Soroka, *Politicization and Polarization in COVID-19 News Coverage*, 42 SCI. COMM. 679 (2020).

99. Wolfgang Strobe et al., *Politicization of COVID-19 Health-Protective Behaviors in the United States: Longitudinal and Cross-National Evidence*, 17 PLOS ONE e0263100 (2021).

100. *E.g.*, Kalinga Seneviratne, *International Media Amidst a Pandemic: COVID-19 Conspiracies, Racism, and Politicization*, *in* COVID-19, RACISM AND POLITICIZATION 8 (Kalinga Seneviratne and Sundeep R. Muppidi, eds., 2021).

101. Cali Curley, Nicky Harrison & Peter Federman, *Comparing Motivations for Including Enforcement in US COVID-19 State Executive Orders*, 23 J. COMP. POL'Y ANAL.: RES. & PRAC. 191 (2021).

102. *E.g.*, Lucy Wang Halpern, *The Politicization of COVID-19*, 120 AM. J. NURSING 19 (2020) (describing various ways in which the response to COVID-19 was politicized and ways differing political parties approached pandemic response).

change data to drum up support for the plan to reopen."[103] Later, Jones became the subject of a criminal complaint for allegedly using internal government communication without authorization, sending a message to workers in the Department of Health that "It's time to speak up before another 17,000 people are dead. You know this is wrong. You don't have to be part of this. Be a hero. Speak out before it's too late."[104] Jones was fired after the state stopped listing deaths from COVID-19 reported by medical examiners, which would have raised death counts substantially above Florida's published record.[105] When her home was raided by armed police and she was later arrested, Jones claimed this was retaliation for whistleblowing.[106]

In Georgia, a dashboard created by public health officials was replaced with one prepared by a contractor with ties to the governor's office. The dashboard showed, for example, a graphic that appeared to document dramatic declines in case counts over time by mis-ordering dates on the axis rather than presenting dates in chronological order.[107] The dashboard also updated its

103. Doha Madani, *Florida Data Scientist in Battle with State over Covid Dashboard Turns Herself In*, NBC NEWS (Jan. 17, 2021), https://www.nbcnews.com/news/us-news/florida-data-scientist-battle-state-over-covid-dashboard-plans-turn-n1254544 [https://perma.cc/8VXD-JGRQ].

104. NSF, *Criminal Case Against COVID-19 Whistleblower Rebekah Jones to Continue*, ORLANDO WEEKLY (June 30, 2021, 10:35 AM), https://www.orlandoweekly.com/news/criminal-case-against-covid-19-whistleblower-rebekah-jones-to-continue-29584117 [https://perma.cc/9F2H-MDC3]; Ava Loomar, *Is Florida COVID Data Activist Rebekah Jones Grifter or Hero, Villain or Role Model, Liar or Whistleblower? Maybe All of the Above*, ORLANDO WEEKLY (June 16, 2021, 4:00 AM), https://www.orlandoweekly.com/orlando/is-florida-covid-data-activist-rebekah-jones-grifter-or-hero-villain-or-role-model-liar-or-whistleblower-maybe-all-of-the-above/Content?oid=29505424 [https://perma.cc/458E-TG8S].

105. Madani, *supra* note 103.

106. *Id.* As of this writing, Jones still faces criminal charges but is running for Congress. Jim Little, *Democrat Rebekah Jones Says She Can Win in Northwest Florida*, PENSACOLA NEWS J. (Feb 8, 2022), https://www.pnj.com/story/news/politics/elections/2022/02/08/rebekah-jones-says-she-can-win-northwest-florida-congressional-seat/9253414002/ [https://perma.cc/VX57-S6AA].

107. Keren Landman, *Behind Georgia's COVID-19 Dashboard Disaster*, ATLANTA (Oct. 24, 2020), https://www.atlantamagazine.com/

coloring scheme (showing the case rate) in a way that changed from day to day to graphically misrepresent the outbreak's extent: side by side graphs from less than a month apart showed identical color schemes despite a near 50% increase in cases.[108] A local media organization used open record requests to examine the creation of a problematic dashboard, finding that:

> the health department had limited input into and no real oversight over the dashboard during its creation and in the months after its launch. Additionally, the sidelining of the health department allowed for errors in the analysis, interpretation, and visualization of the state's COVID-19 data, while simultaneously costing the state tens of thousands of dollars[109]

When the health department attempted to add a lengthy Frequently Asked Questions list regarding the data to the dashboard, the governor's office then removed most of the explanatory content.[110]

---

great-reads/behind-georgias-covid-19-dashboard-disaster/ [https://perma.cc/PB9U-XUCP]; Willoughby Mariano & J. Scott Trubey, '*It's Just Cuckoo': State's Latest Data Mishap Causes Critics to Cry Foul*, ATLANTA J. CONST. (May 19, 2020), https://www.ajc.com/news/state--regional-govt--politics/just-cuckoo-state-latest-data-mishap-causes-critics-cry-foul/182PpUvUX9XEF8vO11NVGO/ [https://perma.cc/5L9T-9ACZ].

108. Landman, *supra* note 107; Nicole Carr, *Public Health Experts Call Out Confusing COVID-19 Data Maps; DPH Set to Make Changes*, WSB-TV ATLANTA (July 23, 2020, 11:30 PM), https://www.wsbtv.com/news/georgia/public-health-experts-call-out-confusing-covid-19-data-maps-dph-set-make-changes/KZPBOLBG2BG2TAMI2N4CQGWESA/ [https://perma.cc/7RL9-SC85].

109. Landman, *supra* note 107 (describing the governor's office outsourcing of dashboard creation to a contractor without general public health specialization, causing a variety of challenges, such as that it contained "inappropriate terminology and lacked sufficient explanatory text; in others, key metrics and tables were absent, or existed where they didn't belong; the graph showing the daily case count did not use shading to indicate a 14-day "pending period" to account for the lag time between a person's onset of symptoms and the confirmation of their positive test result by the state).

110. *Id.*

In New York, former governor Andrew Cuomo faced criticism that his administration had falsified COVID-19 related data on deaths in nursing homes.[111] For a five-month period, officials in his administration allegedly stopped state health officials from reporting the true number of deaths, in order to avoid the data being used for political attacks from the Trump administration.[112] The state-reported tally of deaths included only those nursing home residents who had died inside their nursing home, rather than at, e.g., a hospital. Other states counted those who died at other locations, which made New York's death rate among nursing home residents appear relatively low by comparison.[113] According to reporters, an aide to the governor rewrote the report from its own health department before releasing it, and aides worked actively to prevent the state's public health officials from releasing accurate information to the public.[114]

3. COMMUNICATING PANDEMIC DATA: DASHBOARD STAKEHOLDERS

With the background of public demand for COVID-19 data and its surrounding regulatory landscape in place, together with the development of business intelligence services and the potential politicization of the use of those services, this Section describes a national questionnaire which asked dashboard stakeholders to reflect on those and a variety of related issues. These issues included the challenges respondents faced in creating or managing dashboards, how they perceived the accuracy of data, whether they had been subject to political pressure when creating data visualizations, and so on.

111. *See, e.g.*, Michael Gold & Ed Shanahan, *What We Know About Cuomo's Nursing Home Scandal*, N.Y. TIMES (Mar. 5, 2021), https://www.nytimes.com/article/andrew-cuomo-nursing-home-deaths.html [https://perma.cc/98MD-VAG5].

112. *Id.*

113. *Id.*

114. *Id.*; J. David Goodman, Jesse McKinley & Danny Hakim, *Cuomo Aides Spent Months Hiding Nursing Home Death Toll*, N.Y. TIMES (Apr. 28, 2021), https://www.nytimes.com/2021/04/28/nyregion/cuomo-aides-nursing-home-deaths.html [https://perma.cc/CF38-GW8Q].

A. Questionnaire and Interview Methodology

We gathered contact information (email addresses) for individuals involved in COVID-19 visualization and communication across higher education, K-12 education, state government, and media. This information was obtained from school district websites, directories of state workers, media organization contact information, and so on. [115] Of the 812 individuals contacted, 44 started the questionnaire, with an 86% completion rate, resulting in 38 responses across 18 questions. Six respondents consented to in-depth interviews expounding on their answers to the questionnaire and interviews averaged approximately half an hour. Respondents to the questionnaire and the interview requests included communication officials, advisors to government entities, educational leaders, and COVID-19 dashboard creators. The questionnaire asked for information about the following topics, and the complete questionnaire is included in the Appendix:

- Level of involvement with dashboards or other data visualizations related to COVID-19;
- Sources relied on for statistical data about COVID-19;
- Quality of visual communication about COVID-19 available to the general public;
- The effect of legal regulations and the role of political influence on efforts to communicate pandemic-related data; and
- Ideas for preparing for communication of critical data in future crises.

115. Higher education contacts were obtained by randomly sampling from within the U.S. Department of Education's Database of Accredited Postsecondary Institutions and Programs at https://ope.ed.gov/dapip/#/home [https://perma.cc/Z3NF-5KFN]. School district contacts were obtained by randomly sampling from school districts with at least 20 schools from The National Center for Education Statistics at nces.ed.gov/ccd/pubagency.asp [https://perma.cc/7ABP-DWL9]. An official from each state was included in the sampling list, together with reporters from a variety of large media organizations across the United States. Obtaining email contact information of officials responsible for COVID-19 visualization was a challenging task, because often entity dashboards or other websites did not provide contact information to any specific official responsible for creating or maintaining the dashboard.

B. Results: The Views of Dashboard Stakeholders Across the United States

In this subsection, we summarize the survey comments and interview responses grouped in several general categories: what was most and least effective about COVID-19 data communication, what challenges communicators faced, whether and how data were politicized, the role of public-private partnerships, and so on.[116]

Survey respondents and interviewees reported several aspects of COVID-19 data communication that they found to be particularly effective. Respondents noted that data that was communicated using standardized metrics was the most useful, as it allowed comparisons between and among reporting entities like states, counties, schools districts and higher education institutions. Survey respondents also mentioned that regular and timely updates (especially daily updates) allowed for more impactful decision-making. Visualization formats that provided key information, trends, indicators and "at-a-glance" metrics were noted as being especially useful.[117] At least one person interviewed mentioned that the color schemes used in the dashboards they relied on were helpful, particularly dashboards with dark backgrounds and bright colors to make key data stand out. Several interviewees mentioned that interactive visualizations, rather than static data displays, were critical to them as decision-makers and as mediators of information to their stakeholders. One interviewee noted that the best dashboards look at different data over time.

Survey respondents and interviewees also reported several aspects of COVID-19 data communication that they found to be less effective. Survey respondents commented that conflicting and contradictory information from different authoritative entities (e.g., differences between state and county health department data) caused confusion and made communicating with stakeholders very difficult. Other respondents noted that delays in data updates (anything less than daily) at levels above their jurisdictions resulted in local dashboard data appearing to be

116. We combine answers from the questionnaires, surveys, literature on COVID-19, and so on.

117. We reflect this in our measure of "quality" for dashboard analysis, which looked for the presence of graphics or key indicators. *See infra* note 149 and accompanying text.

inaccurate to end users. The phenomenon of contradictory data resulting from less-than-regular data updates led to real confusion among stakeholders, potentially negatively affecting the level of trust in local data providers. Survey respondents and interviewees identified several other aspects of less effective dashboards, including small fonts, linking out to other dashboards, overwhelming amounts of information, and constantly changing methodologies.[118] On the topic of "referatory-type dashboards," one interview respondent explained that "a home page with lots of links to several dashboards wasn't as useful as a landing page that immediately gave you the data you need."

Communicators faced a variety of challenges. For many entities, such as school districts, the pandemic represented the first time they were asked to communicate data in real-time to stakeholders. Without significant past experience to draw on, dashboards could be difficult to create and manage. One respondent noted they were "building the plane as we were flying." Many worked without templates,constantly modified their work due to the complexity of the data, and simultaneously responded to other COVID-19 related issues. Messaging from public health agencies, such as the CDC and state agencies, could be confusing, methodologies to apply to the data changed, and requirements from public health agencies evolved over time.[119] When asking health agencies for assistance, staffing limitations meant that ability to help was limited. Even obtaining up-to-date data was difficult, as entities searched for qualified individuals to review the data, chose which to display, and managed individual-based data. Creating depictions of data that were complete yet still simple enough to be understood was challenging, as was the difficulty of avoiding logical fallacies within the data.

We explored the perception of whether states or other political entities influenced data reporting through a series of additional questions.[120] First, we asked whether the respondents' efforts to communicate accurate, reliable, and complete

118. The discussions from respondents thus mirrors the general issues found in COVID-19 dashboard design. *See* Wernimont, *supra* note 13.

119. *See, e.g.*, *supra* Section 2B (describing the swiftly changing nature of COVID-19 data reporting requirements from the federal government).

120. *See supra* Section 2D for very visible examples of this pressure.

information about COVID-19 were "influenced" by political interests. Of those responding to this question, nearly half (47%) responded yes. When asked to follow up, respondents noted that they were influenced by political entities outside of public health. One example was navigating a situation where the county health department gave information that contradicted information put out by the governor's office. There was pressure to have the type of data and metrics that supported reopening, and many members of the public made accusations about the data being manipulated for political reasons, whether or not such claims were true. At the same time, most notes about politicization referenced attempts to politically influence decision-making rather than the representation of data itself to the public.

In addition to the question of whether data visualization had been politicized, many respondents reflected on the related issue of whether *responses* to the data were politically influenced. The answer appears to be overwhelmingly yes, and from across the political spectrum. Anecdotally, the presentation of COVID-19 data created confusion and anger in parents, both from conservatives and liberals.[121] Local political issues combined with confusing and changing messaging from authorities to complicate the challenges of responding to a pandemic. One particularly challenging issue was the non-uniform response to new rules and regulations, with, e.g., one school district following local health department guidelines and another not. There might also be little political willpower for following through on previously announced, data-driven decisions. If metrics failed to support a continued policy path, such as reopening schools or removing mask mandates, there was little consequence for ignoring previously established policy. The nature of the data collected changed over time, as it became more difficult to track infection rates with surging caseloads.

To quantitatively assess perception of the general state of COVID-19 data communication and preparedness, we asked respondents three questions. On a seven point Likert scale, respondents were asked (1) "In general, visual communication of information about COVID-19 has been effective in encouraging

121. *Cf.* Jiyoung Han, Meeyoung Cha & Wonjae Lee, *Anger Contributes to the Spread of COVID-19 Misinformation*, MISINFORMATION REV. (Sept. 17, 2020), https://misinforeview.hks.harvard.edu/article/anger-contributes-to-the-spread-of-covid-19-misinformation/ [https://perma.cc/J2G9-JPM7].

public behavior that limited the spread of COVID-19"; (2) "In general, data used in COVID-19 visualizations has accurately reflected infection rates, hospitalizations, and other key statistics."; and (3) "Entities across the United States are prepared to communicate information about future public health crises in reliable and effective ways."[122] The average rating for question (1) was 5.0 on the 1 to 7 scale, suggesting that professionals involved in COVID-19 dashboards believe visual communication has been at least somewhat effective in encouraging behavior that helped limit the spread of COVID-19. For question (2), the average response was higher, at 5.9, suggesting that respondents believed that visualizations were in fact conveying accurate representations of key COVID-19 related statistics.[123] The response to question (3) was lowest, an average of 4.8, indicating a lower belief in preparedness for future, similar communication of pandemic-related data.

While we were impressed with the ability of content creators without extensive technical backgrounds to create sophisticated and useful dashboards, much public communication of pandemic-related data was accomplished through public-private partnerships, or use of contracting entities.[124] This was particularly true for entities without significant resources. For example, higher education institutions often encompass wide-ranging technical expertise in both epidemiology, statistics, and data visualization. Specialized state agencies may have expertise in one area, such as epidemiology, but not others, such as interactive data visualization techniques. Public health agencies without extensive internal expertise often turned to data

122. The Likert scale "is a five (or seven) point scale which is used to allow the individual to express how much they agree or disagree with a particular statement . . . [They] have the advantage that they do not expect a simple yes / no answer from the respondent, but rather allow for degrees of opinion, and even no opinion at all." Saul McLeod, *Likert Scale Definition, Examples and Analysis*, SIMPLY PSYCHOLOGY (2019), https://www.simplypsychology.org/likert-scale.html [https://perma.cc/6EL9-6ZWN].

123. This is welcoming, as in general for dashboards "transparency is an illusion when the data is bad." Lauzardo, *supra* note 12.

124. While not a public-private partnership, the public-public partnership of county health departments with public school districts appears to have played an important role.

visualization specialists through contracts with existing or newly created private entities to create or improve their dashboards.[125]

Perhaps depending on the level of technical expertise within the health department, school district, or other entity, respondents found public-private partnerships helpful on a broad scale from "crucial" to "not at all." On the "crucial" side, private partnerships were described as "indispensable" or taking a "major" role in data visualizations, such in obtaining data, while others noted that private partnerships played a very little role, with most requirements handled internally.[126]

4. COMMUNICATING PANDEMIC DATA: SURVEYING DASHBOARDS NATIONWIDE

This Section describes the methodology and results of a national survey of COVID-19 data dashboards from thousands of entities responsible for COVID-19 communication.

A. Methodology for Studying COVID-19 Dashboards Nationwide

To gain a broad sense of how pandemic-related data were communicated, we studied four groups of entities which often assumed responsibility for creating COVID-19 dashboards. These entities were a mixture of public organizations (states and counties), and organizations with a significant private presence (school districts and higher education institutions).[127] We examined each state's official dashboard and then randomly

125. For example, the authors were hired by the Idaho Department of Health and Welfare in the summer of 2020 to improve the state's official COVID-19 dashboard because of lack of specialized expertise within the department related to certain data visualization techniques. Letter of Notation from Idaho Department of Health and Welfare to Redfish Metrics, LLC (July 14, 2020) (on file with the authors).

126. For instance, if an entity such as a state board of public health had the resources of a GIS office, expertise with ArcGIS would be readily available, and personnel would be capable of producing a dashboard without partnering with consultants.

127. Samples were gathered during 2021 and early 2022, which represents a time period in which entities had more than a year to prepare a dashboard, but still faced ongoing challenges related to COVID-19. Dashboards during this period were also beginning to reflect vaccination rates in addition to COVID-19 case statistics.

sampled a substantial portion of counties,[128] school district, and higher education institutions within the United States to look for the presence of dashboards.[129] In these data, we use the term

128. According to CENSUS.GOV, "The United States total includes 3,006 counties; 14 boroughs and 11 census areas in Alaska; the District of Columbia; 64 parishes in Louisiana; Baltimore city, Maryland; St. Louis city, Missouri; that part of Yellowstone National Park in Montana; Carson City, Nevada; and 41 independent cities in Virginia." U.S. CENSUS BUREAU, STATES, COUNTIES, AND STATISTICALLY EQUIVALENT ENTITIES (last visited Oct. 27, 2021), https://www2.census.gov/geo/pdfs/reference/GARM/Ch4GARM.pdf [https://perma.cc/BYQ7-FTL8].

129. County information was obtained from the United States Census, *Annual Estimates of the Resident Population for Counties in the United States: April 1, 2010 to July 1, 2019*, CENSUS.GOV (Oct. 8, 2021). The universe of higher-education institutions was obtained from accreditation records at U.S. Department of Education's Database of Accredited Postsecondary Institutions and Programs at https://ope.ed.gov/dapip/#/home [https://perma.cc/Z3NF-5KFN]. The sample was drawn from those institutions listed with information under "ParentName." Enrollment numbers for these institutions were gathered from information reported on institution websites during 2021 and early 2022. School districts were sampled from information described in *supra* note 115. Due to the quantity of very small districts less likely to have resources sufficient to create or maintain dashboards, a stratified sample was used which focused heavily on larger school districts. We sampled districts of at least 20 schools at an approximately 90% rate, and districts of less than 20 schools at approximately 2% rate. The statistical information in Tables 1 and 2 and the regressions in Table 6 represent weighted results. *See* Mohammad Ali Mansournia & Douglas G Altman, *Inverse Probability Weighting*, BMJ (2016), https://www.bmj.com/content/bmj/352/bmj.i189.full.pdf [https://perma.cc/X9ZH-Q5A7] (describing inverse population weighting). Unweighted results in Table 6 are similar, with the main exceptions that (1) the effect of having a county dashboard is no longer a significant predictor of the presence of a school district dashboard, and (2) the coefficient on size is small and only marginally significant. This suggests that the effect of having a county dashboard on the choice to create a district dashboard differs between small and large school districts. For each entity, a research assistant examined the entity's contact information and attempted to locate dashboard information related to COVID-19 online. If after a reasonable examination, no dashboard information was located, the entity was coded as not having a dashboard. Thus, so far as dashboard information was provided in an inaccessible manner, it is possible that the related numbers are a slight undercount of the existence of dashboards. Dashboard characteristics were noted from the webpage or potentially from

"dashboard" loosely, in the sense that the institution posted *some* kind of statistical, tabular, or graphical information related to COVID-19.[130] Screenshots of each of these dashboards are available at links in the note below, saved in a series of slide decks which are available online.[131] Table 1 shows the number of sampled entities and summary statistics for each type of entity.

recorded screenshots taken of the webpages. Many school districts in these data did not contain links to webpages, and so internet searches were performed to find their online information. Data on school districts with common names which did not have URLs in the survey data were checked for location in the appropriate state. In a handful of rare cases, school districts with identical names and sizes across states — for which it was impossible to verify sampling — were omitted.

130. If a school district only provided a link to a *state-created* dashboard or a county dashboard, we did not count this as a school district dashboard. The classification of "dashboard" or not relied on making somewhat fine-grained distinctions between entities, which often relied on the discretion of the coder (for instance, whether text reporting numbers without tabular form but referred to by the entity as a "dashboard" constituted a dashboard for our purposes). For the many entities that simply linked to Google sheets with tabular information, see *infra* note 145. Because of this, we make the distinction below between this classification and dashboards that contained graphics or key indicators set off in large fonts, which are indicative of effort to summarize or aggregate data in truer "dashboard form", and we make the collection of dashboards available for alternative classifications. *See infra* note 131.

131. Dashboards for states, counties, higher education institutions, and school districts can be accessed at https://redfishmetrics.com/covid-dashboards [https://perma.cc/4A3G-WT2W]. Some screenshots were not counted as "dashboards" in the quantitative coding but were retained for comparison purposes. *See infra* note 145.

Table 1: Summary Statistics of Dashboards in the United States

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Entity	**Final Sample Size**[132]	**Sample Average Entity Size**	**Total Dashboards Collected**	**Probability of Providing a Dashboard**	**Probability of Graphs or Key Indicators**	**Most Common Software**
States	51	6,499,006 (population)[133]	51	100.0%	88.2%	Tableau
Counties	1,000	94,700 (population)	279	27.9%	22.1%	ArcGIS
School Districts	981[134]	4.97 (schools)[135]	504	61.0%[136]	30.1%	Basic HTML
Higher Ed	901	5,922 (enrollment)	320	35.5%	19.6%	Basic HTML

132. Some school districts and higher education institutions draw from the sources described in *supra* notes 115 and 129 were closed, did not provide information about size of the institution online, and so forth. The final sample size reflects the number of institutions with information found publicly available.

133. State populations were drawn from the 2020 United States Census results at https://data.ers.usda.gov/reports.aspx?ID=17827 [https://perma.cc/7U8Y-NEVD]. This number includes the 50 states and the District of Columbia.

134. As noted in *infra* note 154, the final reported sample size for school districts and higher education institutions reflects that some entities listed in databases had closed, otherwise had no information available online, and so on.

135. Enrollment in school districts was not available, and so we use number of schools within the district for a measure of entity size.

136. This is a weighted probability, per *supra* note 129. The unweighted average is 504/981 or 51.4%.

Totals/ Averages	*2,933*	*NA*[137]	*1,154*	*42.6%*[138]	*25.2%*	*Basic HTML*

Notes: In Column (4), figures reflect total dashboards collected which published online information about the size of the entity. Additional dashboards for entities which did not provide this information online are saved in the collections of dashboards. In Column (7), Basic HTML was the most common software used across all entities.

Overall, we examined almost three thousand entities across the United States that would potentially be expected to provide a COVID-19 dashboard. Among these entities, we found over 1,200 dashboards, of widely varying quality, format, software, and graphical design. Each state created a COVID-19 dashboard, and then among other entities, approximately a third of each (between 27 and 38 percent) created a dashboard. Figure 1, below, shows the total number of entities examined for the presence of a dashboard in each state, including the state itself, higher education institutions, counties, and school districts. The depth of color indicates the number of sampled entities.[139] Figure 2 then expresses the proportion of those sampled entities from which a dashboard was collected. On the low end, only 12% of entities in South Dakota produced a dashboard, while on the high end, 74% of entities in Washington state produced a dashboard.[140] Figure 3 then shows the proportion of entities within a state that contain dashboards with either graphics or key indicators—data visualizations which would easily be recognized as "dashboards" rather than mere tabulations of data.

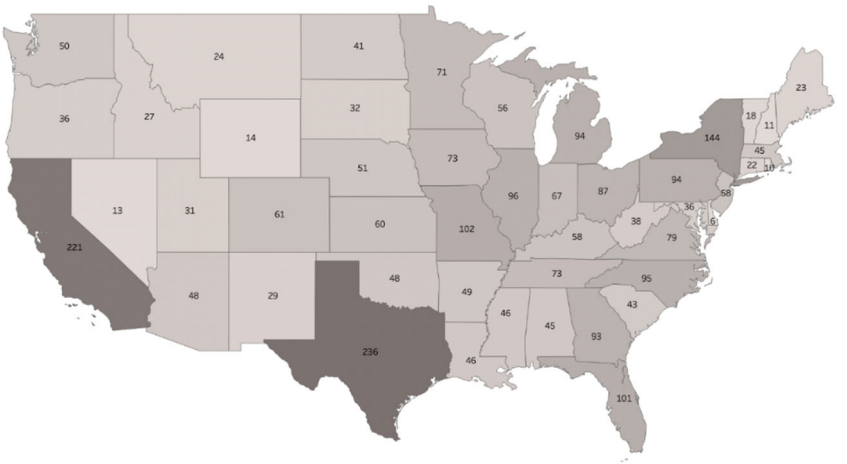
137. As the number of schools (districts), enrollment (higher education institutions) and population (states and counties) each represent different metrics, an average "size" figure is not given.

138. This represents a weighted average of the probability of creating a dashboard across all entities. Column (6) is calculated similarly.

139. Our sampling procedure was not stratified by state, hence more populous states had more entities examined.

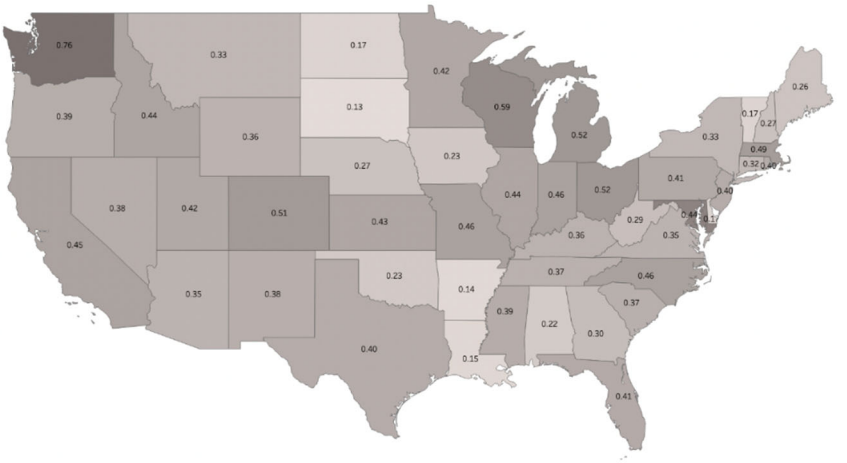
140. While the sample size for some states is low, there appears to be little correlation between region of the country and the proportion of entities creating dashboards. Percentages were calculated as simple aversages of the 0/1 variable indicating presence of a dashboard.

Figure 1: Number of Sampled Entities, by State



Notes: Color scale indicates the number of sampled entities (states, counties, higher education institutions, and school districts) within the state. States and territories not shown are Alaska (17 entities), Hawaii (6 entities), and the District of Columbia (9 entities).

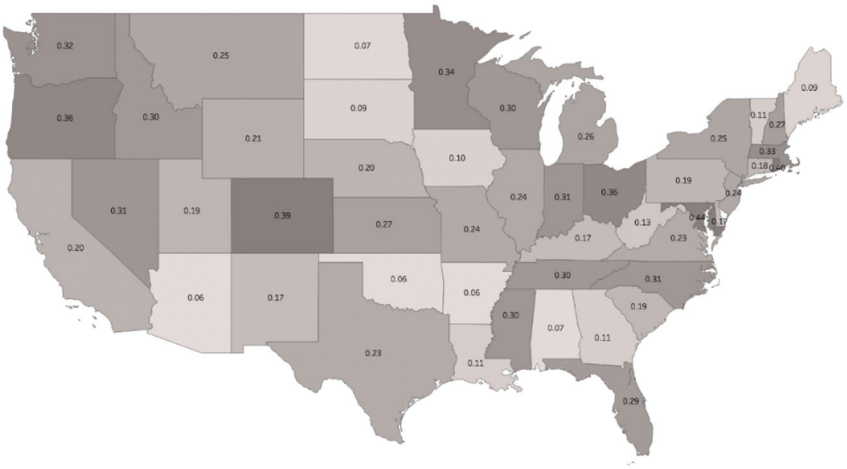
Figure 2: Proportion of Entities with a COVID-19 Dashboard, by State



Notes: Color scale indicates the proportion of sampled entities (states, counties, higher education institutions, and school

districts) publishing a COVID-19 dashboard as defined in the text. Proportions not shown are Alaska (0.47), Hawaii (0.33), and the District of Columbia (0.44).

Figure 3: Proportion of Entities with a COVID-19 Dashboard that Contains Graphics or Key Indicators, by State



Notes: Color scale indicates the proportion of sampled entities (states, counties, higher education institutions, and school districts) publishing a COVID-19 dashboard with graphics or key indicators set apart in large fonts. Proportions not shown are Alaska (0.35), Hawaii (0.17), and the District of Columbia (0.44).

Next, as noted above, we categorized the dashboards in several fundamental ways. First, we categorized according to which software was used to create the dashboards. We categorized the software used according to the main business intelligence entities described in Section 2 above and a variety of additional options: "Basic HTML," (meaning some form of statistical rap information presented in a webpage without apparent use of a business intelligence service), Tableau,

PowerBI, ArcGIS, Smartsheets,[141] Highcharts,[142] Infogram,[143] Google Data Studio,[144] other Google products,[145] or Other / Unknown.[146] We also assessed dashboards according to measures of quality. There are extensive opinions about best practices in communicating data in dashboard form.[147] These include suggestions such as export capability, interactivity, and graphical representations of data.[148] Here, we categorize dashboards

141. *See Welcome to Smartsheet*, SMARTSHEET.COM (2022), https://www.smartsheet.com/ [https://perma.cc/TSD8-NGEV] ("The dynamic workspace that moves your business forward.").

142. *See Simply Visualize*, HIGHCHARTS.COM (2022), https://www.highcharts.com/ [https://perma.cc/A9BN-VYKQ] ("Rock-solid and incredibly flexible charting library made for developers.").

143. *See Create Engaging Infographics and Reports in Minutes*, INFOGRAM.COM (2022), https://infogram.com/ [https://perma.cc/M87Y-D6NL] ("Infogram is an intuitive visualization tool that empowers people and teams to create beautiful content.").

144. *See Welcome to Data Studio*, GOOGLE.COM (2022), https://support.google.com/datastudio/answer/6283323?hl=en [https://perma.cc/P5XZ-QLTF] ("Data Studio is a free tool that turns your data into informative, easy to read, easy to share, and fully customizable dashboards and reports.").

145. These would include Google Sheets or Google Docs with dashboard-like information made publicly available. Many entities published what essentially amounted to raw COVID-19 data in Google Sheet form with no attempt to aggregate into more usable form or convert to HTML. The line between calling these "dashboards" or not was difficult to draw. As a straightforward delineation in this category, we counted use of a Google office product without any graphical information or key indicators as "not a dashboard" for our purposes. We have retained some of these for example purposes in the slide decks in *supra* note 131.

146. These include, e.g., Airtable (https://www.airtable.com/), Datawrapper (https://www.datawrapper.de/), and Clear Impact (https://clearimpact.com/).

147. *See, e.g.*, Mahtab Karami, Mostafa Langarizadeh & Mansoor Fatehi, *Evaluation of Effective Dashboards: Key Concepts and Criteria*, 11 OPEN MED. INFO. J. 52 (2017).

148. *See, e.g.*, Naleef Fareed, Christine M. Swoboda, Sarah Chen, Evelyn Potter, Danny T. Y. Wu, Cynthia J. Sieck, *U.S. COVID-19 State Government Public Dashboards: An Expert Review*, 12 APPL. CLIN. INFOMATICS 208 (2021) (examining features across state dashboards).

according to two accessible measures of quality: whether the dashboard had graphics, or not, and whether the dashboard summarized key indicators in a font that stood as larger than surrounding text. These represent two simple and relatively objective assessments of dashboard quality that distinguish from entities that simply reported tables of numbers and those that attempted to aggregate the data for their stakeholders in an easy-to-read manner.[149]

B. Results: The Role of Resources and Software in Dashboard Creation

This survey of dashboard creation and quality across the United States provides multiple lessons for the visualization of pandemic-related data. We highlight three. First, the survey provides insight into the relationship between available resources and the ability to create a dashboard that scores well on quality metrics. Our primary ability to examine resources comes through data gathered on the size of entities, meaning, e.g., county population or number of schools in a school district. Second, closely related to the idea of resources, is the role of software across institution type and whether that software enabled the creation of certain dashboard elements. Third, we offer qualitative anecdotal evidence of best practices across the range of dashboards collected from around the country.

First, we consider the question of resources. Creating a public-facing dashboard of any quality requires institutional resources committed to the task, including expertise or the ability to acquire expertise, in software. For a higher quality dashboard, such as one involving maps or graphics, additional resources are required, such as the ability to utilize GIS software. Once created, maintaining a dashboard requires further commitment of resources, from public health officials willing to provide data to software operators with time to update visualizations.[150] The best

149. There is an incredible variety of ways these data were presented, and so classifying dashboard elements into dichotomous categories is a challenging task, and other classifications may be equally valid, such as counting bold fonts as quality key indicators, rather than fonts set apart based on size, and so on. We have categorized these data in one way, but future research with differing categorizations may be valid.

150. For example, *see* Hosam Elattar, *Saddleback Valley School District Angers Parents by Ending COVID Dashboard During Surge*, VOICE OF OC (Jan. 14, 2022), https://voiceofoc.org/2022/01/saddleback-

available proxy for resource potential that is comparable across a variety of entities in the United States is simply the *size* of the entity in terms of population. For counties, population is a strong determinate of available resources.[151] Similarly for higher education institutions, enrollment is a general predictor of resources, because the larger the enrollment base, the more likely the institution is to have departments and staff with expertise in the areas needed to develop and maintain dashboards.[152] For school districts, we lack data on direct enrollment of students, but we do observe the number of schools governed by the district. This is an imperfect proxy, because schools themselves may vary in size, but—as will be shown below—examining school number and the probability of creating a dashboard reveals patterns similar to those found by looking at population and enrollment for counties and institutions of higher education.

Table 2 shows the relationship between size of entity and the probability of creating a dashboard, as well as the probability of

valley-school-district-angers-parents-by-ending-covid-dashboard-during-surge/ [https://perma.cc/GP88-JBYY] (quoting a district website as noting "With a focus on our students, we have been funneling our District Office personnel resources to our school sites so that in-person student learning is not interrupted as we experience isolation periods of school staff members.").

151. For example, in one state with information available, Oklahoma, average total expenditures in 2015 were more than twice as high for counties with more than 45,000 residents (~$7.2M) compared to counties with 20,000-45,000 residents (~$3M) and more than three times as high compared to counties with fewer than 20,000 residents (~$2.2M). OKLAHOMA STATE, FACT SHEETS (2015), https://extension.okstate.edu/fact-sheets/images/comparison-of-county-government-finances-by-size-groupings/table-3-3.pdf [https://perma.cc/7Z4Y-ZG26].

152. In Idaho, for example, as in other states, legislative appropriations for public higher education institutions are based primarily on institution size, as determined by student enrollment numbers. In 2021, Boise State University had an enrollment of 24,000 students and received $255M, Idaho State University had an enrollment of just under 11,700 students and received $224M, University of Idaho had an enrollment of 10,700 students and received $193M, and Lewis-Clark State College had an enrollment of almost 4,000 students and received $53M. FY2022 Legislative Budget Book, 1-055 Colleges and Universities. IDAHO LEGIS. SERV.'S OFF., IDAHO LEGIS. BUDGET BOOK FOR FISCAL YEAR 2022 (2021), https://legislature.idaho.gov/lso/bpa/pubs/lbb/?y=2021 [https://perma.cc/K8QF-XVLG].

creating a dashboard that contains graphical elements and key indicators. For each entity type, the table shows information divided at the 25th, 50th, and 75th percentiles of the size of the institution. For example, for higher education institutions, the table shows summary statistics for schools in the bottom 25 percent of enrollment size. For these smallest higher education institutions with enrollment of less than about 580 students, the probability of creating a dashboard is close to zero. For institutions between the 25th and 50th percentiles in terms of enrollment, the probability increases to approximately 31%. This probability then increases again for institutions between the 50th and 75th percentiles, to almost 46%. Finally, the probability for the largest institutions, those with enrollments of over about 6,700 students, the probability of having a dashboard is over 64%.

Other than states, which universally had some form of a dashboard, the same general pattern between entity size and dashboard probability exists. For counties, we surveyed information from 1,000 randomly sampled counties across the United States, representing about a third of all counties in the United States. Of those thousand counties, 279, or about 28%, had a dashboard of some sort. Counties in the lowest quantile of population were very unlikely to have a dashboard, with a probability of only six percent. This jumped substantially for counties between the 25th and 50th percentiles of population, which had an almost 30% chance of creating a dashboard. This rises to almost 50% for counties between the 50th and 75th percentiles. The most populous counties, above the 75th percentile, were more likely to have a dashboard than not, with a probability of almost 55%.[153]

Our sample of school districts also shows this pattern. From our sampling, we obtained information on 981 districts from a comprehensive list of the approximately 19,000 districts across the United States, focusing on districts with at least 20 schools.[154] The overall percentage of districts with a dashboard was higher than that of counties and higher education institutions, at around 60%. Our measure of size—number of schools within a district—

153. In a regression setting, the relationship between log of county population and the existence of a dashboard is not significant. *See infra* Table 6.

154. This is the total sample size after accounting for listed school districts that had, e.g., ceased operations at the time of data collection.

shows a pattern similar to that of counties and higher education institutions. Schools in the bottom 25th percentile had roughly a 23% likelihood of creating a dashboard, which rose to upwards of 66% for larger districts.

While there appears to be a strong and constituent relationship between size of an entity and the probability of creating a dashboard, the relationship between entity size and probability of creating a dashboard with graphics or key indicators is more complex. Column 6 shows these probabilities among the dashboards collected in the sample. First, even states did not universally create dashboards with these elements of quality.[155] While the number of states in each quantile is relatively low, the percentages remain below 100%, and there is little relationship between the size of the state and creation of a "high quality" dashboard. With other entities, however, there is a strong relationship between size of the entity and whether the dashboard they created contained graphics or key indicators.

Table 2: COVID-19 Dashboard Statistics by Quantiles of Entity Size

(1)	(2)	(3)	(4)	(5)	(6)
Entity	**Quantile**	**Size Metric for Quantile Calculation**	**Dashboard Probability**	**Probability of Dashboard with Graphics or Key Indicators**	**N**
State	0–25th	Population	100.0%	92.3%	13
	25–50th		100.0%	92.3%	13
	50–75th		100.0%	75.0%	12
	75–100th		100.0%	92.3%	13
County	0–25th	Population	6.0%	3.2%	250
	25–50th		20.4%	13.6%	250
	50–75th		29.6%	23.2%	250
	75–100th		55.6%	48.4%	250

155. See the collection of state dashboards linked in *supra* note 130 (considering state landing pages gathered in the dashboard collection). It may be that states with dashboard elements across multiple pages did include these elements on some dashboard pages but were not counted in the assessment of elements.

School District	0–25th	Number of Schools	23.1%	4.4%	251
	25–50th		54.7%	26.4%	255
	50–75th		61.6%	26.7%	232
	75–100th		66.7%	37.0%	243
Higher Education	0–25th	Enrollment	0.9%	0.4%	226
	25–50th		31.0%	11.5%	226
	50–75th		46.2%	24.9%	225
	75–100th		64.3%	42.0%	224

Notes: Dividing quantiles (25th, 50th, and 75th) for state populations are 1,816,411, 4,505,836, and 7,428,392. For counties, these are populations of 10,454, 25,960, and 70,628. For school districts in these data, these are 5, 25, and 41 schools. Probabilities for school districts are unweighted. Quantiles for higher education enrollment are 588, 2,300, and 6,908 students. Probabilities for school districts are weighted according to sampling probabilities. See text for definition of the existence of a dashboard.

Second, we consider the role of software. Due to advances in business analytic technology, constructing and deploying a simple dashboard is within the expertise of many individuals without computer science or statistical backgrounds.[156] Many creators borrowed design elements from the successful dashboards they saw from other entities. State agencies, particularly in states with less resources, used dashboards from larger states or other large institutions as starting points for their designs. Implementing those designs required many entities to find software solutions that were cheap, efficient, and possible to implement within their existing in-house technical expertise. Table 3, below, shows the software used to create dashboards in our sample, along with percent of all dashboards in the sample using the software, and the percentage of those dashboards which contained graphics or key indicators. By far the most common source of creating what we would term a dashboard was what appeared to be basic HTML programming without use of a third-party business intelligence service. However, few of these basic HTML dashboards utilized graphics and key indicators. Dashboards created using third-party services were much more likely to contain graphical elements.

156. *See supra* Section 2C (describing the development of business intelligence software).

Table 3: Use of Software in Creation of COVID-19 Dashboards

(1)	(2)	(3)	(4)	(5)
Software	**N**	**Proportion of Total Dashboards**	**Proportion with Graphics**	**Proportion with Graphics or Key Indicators**
Basic HTML	471	39.3%	1.5%	22.4%
Other/ Unknown	299	24.8%	54.8%	70.6%
ArcGIS	101	8.4%	70.3%	96.0%
PowerBI	83	6.9%	74.7%	94.0%
Tableau	81	6.7%	76.5%	88.9%
Google Data Studio	62	5.2%	57.1%	95.2%
Infogram	16	1.7%	55.0%	100.0%
Other Google	11	1.3%	68.8%	93.8%
Highcharts	9	0.7%	77.8%	100.0%
Smartsheets	8	0.7%	37.5%	62.5%

Notes: Results are not stratified by type of entity, and so results are shown unweighted for school district sampling probabilities. For the category of "Other Google," see note 145. For the category of "Other/Unknow," see note 146.

Next, Table 4 uses the quantiles from Table 2 and examines the use of software by entity type. Column (4) shows the most common software used, and Column (5) shows the second most common software used. Small and large states split between ArcGIS for the smaller states, while larger states favored Tableau (but not universally so; e.g., Florida used ArcGIS). Counties showed the same pattern, but with smaller counties using basic HTML and larger counties favoring ArcGIS. School districts tended not to employ software with licensing fees and largely favored basic HTML, and as a second choice generally favored Google Data Studio. Like school districts, higher education institutions also relied on basic HTML, except that they generally relied on PowerBI as a second choice.

Table 4: Software Use by Entity and Quantile of Entity Size

(1)	(2)	(3)	(4)	(5)	(6)
Entity	**Quantile**	**Size Metric**	**Most Common Software**	**Second Most Common Software**	**N**
State	0–25th	Population	ArcGIS	Tableau	11
	25–50th		ArcGIS	Tableau	10
	50–75th		Tableau	ArcGIS	9
	75–100th		Tableau	Basic HTML	10
County	0–25th	Population	Basic HTML	Tableau	11
	25–50th		Basic HTML	ArcGIS	34
	50–75th		ArcGIS	Basic HTML	47
	75–100th		ArcGIS	Basic HTML	112
School District	0–25th	Number of Schools	Basic HTML	Other Google	43
	25–50th		Basic HTML	Google Data Studio	88
	50–75th		Basic HTML	Google Data Studio	96
	75–100th		Basic HTML	PowerBI	116
Higher Education	0–25th	Enrollment	PowerBI	*NA*	1
	25–50th		Basic HTML	PowerBI	62
	50–75th		Basic HTML	PowerBI	81
	75–100th		Basic HTML	PowerBI	111

Notes: Table does not include those dashboards categorized as using Other / Unknown software, and so the figures in Column (6) differ from those in Table 2, Column (6). In Column (6), the figures shows are the number of entities with a dashboard within the quantiles calculated from all institutions.

Third, the process of interacting with dashboard stakeholders from around the country pairs with these results to offer significant lessons. Even a cursory examination of the ways entities attempted to communicate pandemic-related data shows that dashboards varied *enormously* in terms of quality, data reported, use of graphics, presence of interactive elements, the software used, and so on. In general, effective data communication about COVID-19 followed several similar key elements, and provided substantial benefits. Effective communication was timely, particularly when daily updates were possible.[157] It also allowed for quick cognitive processing of key information, showing key indicators and trends at a glance.[158] Graphics, rather than text alone, were able to convey information to the public more effectively and efficiently.[159] Effective communication also allowed for comparison with other entities, such as states not only showing their own trends, but allowing comparison with other states.[160] Particularly when done effectively, visual communication of COVID-19 data helped the public anticipate and plan for the future, enabled stakeholders to make data-driven decisions, and increased trust in government. This final point–increasing trust in government—is particularly important in light of growing political polarization in the United States. A basic level of trust in government, from the national to the local level, will be critical in responding effectively and collectively to future crises. Finally, good data visualization helped stakeholders understand the magnitude of the challenge of responding to COVID-19. This shared clarity and understanding enabled cooperation between policy-makers and those affected by their decisions.[161]

157. At the same time, dis-uniform update times between various entities caused challenges. For example, if a school district updated its website weekly due to resource constraints, parents observing daily updates from other sources may become distrustful of the district information because of discrepancies caused by misaligned timing of updates.

158. *See supra* note 35 and accompanying text.

159. *Id.*

160. Comparable information is enhanced by providing data in identical or standardized metrics.

161. What many might consider small issues with dashboards could become controversial. One organization, interviewed confidentially, used a yellow to red color scale, but people complained that the

At the same time, many types and methods of communicating pandemic-related data were less effective. Respondents to the questionnaire and interviews noted that various sources of data contradicted others, and that differences in timing between data updates among those sources led data to look inaccurate. Some dashboards had small fonts, pointers to other dashboards that respondents found less effective, and dashboards that were generally overwhelming. If statistical or reporting methodologies changed over time, comparisons in how the disease progressed was difficult. At a higher level, if states interfered with data available to public health departments, or if the data itself was self-reported, ensuring or labeling the data as "accurate" was challenging. Our survey of dashboards across the country confirmed this: many dashboards made no effort to establish trends over time, deviations from a baseline, or to graphically convey the status of the disease. While offering a basic table of COVID-19 numbers might be useful to some, presenting a simple table of information fails to take advantage of the rich modern environment for quickly communicating more useful and actionable information about the local state of public health through modern data visualization techniques.

C. Results: A Regression Analysis of Dashboard Creation

To examine the role of resources together with other factors that may have led to entities creating a dashboard or investing resources in a quality dashboard, we paired the data described above with county-level demographic characteristics from the 2019 American Community Survey (ACS), prior to creation of widespread visualizations about the pandemic.[162] The ACS does not report aggregate data across every county, but a sizeable number of counties overlap between our sampling and ACS coverage, providing an ability to pair dashboard creation with the ACS's deep information on characteristics of communities. For school districts and higher education institutions, we relied on the zip code of the institution's address to pair with county-level

color scale was too grim and represented death. Selecting colors for reported male versus female data could also be problematic, as people associate colors with external concepts.

162. The ACS accompanies the United States Census, but surveys a proportion of the United States in greater depth, and with greater frequency.

data.[163] We also paired these data with information on the presence of a metro area,[164] and data on voting from the 2020 presidential election.[165] Table 5 shows summary statistics for these variables among those counties that paired with counties in our sample. These include the population within the county, median income, the percentage of the population 16 and over working from home, the percentage working in scientific, management, or artistic occupations,[166] the proportion of minority populations within the county,[167] the proportion of the population over age 85, and the percentage of the county with broadband internet access.

163. Zip code matching was performed using zipcodeR. *See Gavin Rozzi*, *Introduction to zipcodeR*, GITHUB (Sept. 22, 2021), https://gavinrozzi.github.io/zipcodeR/articles/zipcodeR.html [https://perma.cc/6ZPY-9QYH]. Some entities failed to list zip codes in their accreditation data or listed zip codes that appeared to not match their state. These were not used in the regressions, although their dashboards are retained in the slides in *supra* note 131 and reported in the summary information in Section 4B which does not rely on county-level matching. In addition, although we relied on parent locations when matching zip codes to counties, matching higher education institutions to a specific county location is a challenging exercise, because higher education institutions may often branch across many disparate locations or have sub-entities of the same parent organization that also branch across locations.

164. U.S. DEP'T AGRIC.: ECON. RES. SERV., *Download Data*, https://www.ers.usda.gov/data-products/county-level-data-sets/download-data.aspx [https://perma.cc/7M9P-2477] (last visited Feb. 14, 2022). Rural classification codes are described at *Rural Classifications: Overview*, U.S. DEP'T AGRIC.: ECON. RES. SERV. (June 17, 2021), https://www.ers.usda.gov/topics/rural-economy-population/rural-classifications/ [https://perma.cc/4RRG-MTJU]. We employed the "metro" versus "non-metro" classifications to distinguish counties.

165. *Data*, MIT ELECTION DATA & SEC. LAB, https://electionlab.mit.edu/data [https://perma.cc/L9P3-9SSE] (last visited Feb. 14, 2022).

166. The proportion of the population age 16 and over employed in management, business, science, and arts occupations.

167. We define this as 100 minus the aggregate self-reported percentage of those within the county "having origins in any of the original peoples of Europe, the Middle East, or North Africa." U.S. CENSUS BUREAU, *About the Topic of Race* (Dec. 3, 2021), https://www.census.gov/topics/population/race/about.html [https://perma.cc/R28A-J7SR].

Table 5: Summary Statistics for County-Level Regression Covariates

(1)	(2)	(3)	(4)	(5)	(6)
Variable	**Observations**	**Mean**	**Std. Dev.**	**Min**	**Max**
Population	257	325,722	451,003	62,385	3,338,329
Median income	257	67,004	19,166	35,035	151,800
Republican vote share	257	0.51	0.15	0.09	0.82
Percent minority	257	22.28	15.38	3.5	84.0
Metro area	257	0.88	0.32	0	1
Tech and arts jobs	257	38.59	7.76	25.9	65.0
Percent home work	257	5.51	2.37	1.2	17.6
Population over 85	257	1.98	0.67	0.8	5.7
Broadband access	257	86.12	5.34	65.3	97.0

Notes: Summary statistics are shown for those counties in which ACS sampling data and data from the sampling described above exist. See Column (1) of Table 6.

These variables attempt to capture many significant features of the location which may affect the creation of dashboards, both from the supply side (e.g., perhaps dashboards are created more easily in areas with technical expertise among the population) and the demand side (e.g., how much of the population in the area is of an age that is considered highly susceptible to serious complications from COVID-19). These variables certainly do not capture all aspects of the decision whether to create a dashboard or not, but they do capture several major areas of consideration. Also, for school districts and higher education institutions, these

represent factors in their area, not the particular expertise available at the districts or institutions themselves.

With those conditions, Table 6 then shows the results of these regressions. Each regression shown includes state fixed effects, meaning that the results shown compare entities to other entities within the same state, controlling for potentially confounding variables that differ at the state level, such as state-level policies related to COVID-19. The first three columns show results for whether an entity was classified as having a dashboard or not in a broad sense,[168] and the final three columns show results when classifying entities as having dashboards when they contained either graphics or key indicators set off in large fonts. When comparing entities to other entities within the same state, many variables which we suspected to be related to the creation or quality of dashboards were not significant, yet the statistics do show a series of intriguing relationships. The size of entity is a significant positive predictor of whether school districts and higher education institutions created a dashboard, and for whether these entities invested in dashboards that contained graphics or key indicators. The presence of a county dashboard is a significant positive predictor of whether school districts within the county created dashboards with graphics or key indicators, meaning that school districts in counties with dashboards were *more* likely to have built a dashboard that went beyond, e.g., merely reporting cases in table form.[169]

168. *See supra* note 129.

169. We do not observe the timing of county versus higher education or school district dashboard construction, and so this analysis represents the relationship between whether a county contained a dashboard when examined during 2021 or early 2022 to whether other entities contained dashboards when examined during 2021 and early 2022.

Table 6: County, School District, and Higher Education Regression Results

Dependent Variable	(1) Dash-board	(2) Dash-board	(3) Dash-board	(4) Graphics /Key Indicators	(5) Graphics /Key Indicators	(6) Graphics /Key Indicators
Sample	Counties	School District	Higher Educ.	Counties	School District	Higher Educ.
Size of Entity	0.053	0.006**	0.109***	0.074*	0.004**	0.084***
	(0.037)	(0.003)	(0.015)	(0.039)	(0.001)	(0.015)
County Dashboard		0.006	0.008		0.353***	-0.103
		(0.242)	(0.099)		(0.126)	(0.096)
Log median income	-0.178	-0.036	0.033	-0.217	0.073	-0.373
	(0.247)	(0.294)	(0.375)	(0.195)	(0.321)	(0.354)
Republican vote share	-0.668	0.522	0.296	-0.789*	0.598	-0.361
	(0.409)	(0.751)	(0.715)	(0.424)	(0.926)	(0.599)
Percent minority	-0.008**	-0.002	0.000	-0.008**	-0.003	-0.002
	(0.003)	(0.007)	(0.005)	(0.004)	(0.006)	(0.004)
Metro area	0.113	0.512	-0.286	0.089	0.249	-0.511***
	(0.129)	(0.364)	(0.176)	(0.143)	(0.174)	(0.165)
Tech and arts jobs	0.008	0.009	-0.008	0.013*	0.011	-0.009
	(0.007)	(0.013)	(0.009)	(0.008)	(0.015)	(0.009)
Percent home work	-0.004	-0.089*	0.016	-0.010	-0.105**	0.036
	(0.014)	(0.046)	(0.028)	(0.013)	(0.043)	(0.031)
Population over 85	0.051	0.012	-0.138*	0.017	-0.013	-0.006
	(0.046)	(0.113)	(0.070)	(0.049)	(0.161)	(0.117)
Broadband access	0.002	0.037	-0.002	-0.002	0.006	0.008
	(0.009)	(0.028)	(0.018)	(0.007)	(0.015)	(0.017)
Observations	257	288	209	257	288	209
R-squared	0.492	0.376	0.417	0.457	0.437	0.380
State FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each column represents a separate regression. The dependent variable in Columns (1) to (3) is a 0/1 indicator for

the presence of a dashboard, as defined in the text. In Columns (4) to (6) the dependent variable is a 0/1 indicator of quality (the presence of graphics OR key indicators). "Size of Entity" is log of county population in Columns (1) and (4), number of schools per school district in Columns (2) and (5), and log enrollment at higher education institutions in Columns (3) and (6). "County Dashboard" is the presence of a county-level dashboard in the county containing the zip code of the entity in columns (2)-(3) and (5)-(6). Other variables are described in text. Sample sizes differ between columns because of varying numbers of entities between columns. Each regression includes state fixed effects, and standard errors are clustered by state. Significance is denoted with *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Demographic characteristics are generally insignificant across the models. Despite significant political differences among approaches to the pandemic, the share of the county voting Republican in the 2020 election is generally not a significant predictor of dashboard creation. The percentage of the population comprised of non-white groups is a significant negative predictor of the probability of creating a dashboard, although the effect is relatively small compared to the other coefficients. Higher education institutions in metro areas appear less likely to create dashboards than those in other areas, perhaps because many higher education institutions in metro areas in our sample are online-focused entities. The share of the population working from home is negatively associated with creating a dashboard for school districts, perhaps indicating lower demand for dashboards in the area. Perhaps surprisingly, the share of the population most susceptible to COVID-19 was generally not a significant predictor of creating a dashboard, nor was broadband internet access, which may serve as a proxy for the ability to access dashboards. This may be because many dashboards were accessed through mobile devices, rather than through broadband home-internet access.

The relationships in Table 6 do not represent the results of an experiment in which, for example, counties were randomly assigned the responsibility of creating a dashboard. Nor do they represent the results of a natural experiment which attempts to use the timing of COVID-19 infection rates or other factors as a source of exogenous variation driving dashboard decisions. Even with multiple variables and state controls in the model, much of

the variation in dashboard creation remains unexplained.[170] Thus, making causal claims from these results would be unwarranted without making very strong assumptions about the data. At the same time, the presence of significant relationships among key variables even after controlling for the state of the entity and a variety of other factors potentially related to the creation of dashboards suggests that, for instance, resources do matter for many institutions considering whether to communicate about COVID-19 using data visualizations, and that perhaps some of those institutions may have considered local conditions in their decision to construct a dashboard. Further study using techniques likely to establish causal relationships may shed additional light on these relationships and provide additional insight.

5. Conclusion: Suggestions for the Future

To conclude, this Section offers specific suggestions for the future based on the survey of dashboards and the national questionnaire and interview process. Suggestions can be broadly categorized into five groups: (1) reducing legal frictions, (2) creating data visualization policies before a crisis, (3) creating policies for key metrics before a crisis, (4) improving data-sharing culture, and (5) increasing resources available for data communication.

First, this study suggests the need to minimize legal frictions that impede the data visualization process. Having policies in place which clear up legal requirements that impact data reporting is helpful. Knowing before a crisis comes what HIPAA does or does not require, what confidentiality requirements imposed by employment law entail, and so on, will remove a barrier towards quickly creating useful visualizations. Perhaps more importantly, data visualization architecture can be created with sufficient flexibility to respond to changes in reporting requirements. For example, data visualizations which are easy to update by public health staff rather than by computer scientists may aid in the ability to respond to changes such as the hospital data reporting alterations described above.[171]

170. The R-Squared values represent the amount of variability in dashboard creation explained by the model. These values never exceed 0.50 across the models in Table 6.

171. *Supra* Section 2B.

Second, aspects of visualization itself can be prepared by having sound visualization policies in place ahead of time. This removes the "build the plane as we're flying it" problem described above. Having a designated single point of contact for reporting data that flows into dashboards may simplify the reporting process, so long as that point of contact has the necessary technical support to manage the data flow. This study suggests it is important to avoid reporting data that are only interesting to a few specialized audiences. Specialized audiences might be better served with *internal* visualizations targeted to those with specific needs and skillsets. Fundamental data that are accurate, relevant to most people, and conveyed through key, pre-planned indicators, appear to be the strongest way to communicate effectively to a wide, general audience.

Third, to better prepare for the technical aspect of communication, it would be helpful to identify which data to collect, how to ensure its reliability, and which key metrics and indicators to report ahead of time. These decisions would be paired with a program in place for where to obtain information on those indicators and how they will be displayed. Deciding which data to collect and key metrics to report ahead of time offers several advantages: It encourages the highlighted presentation of key indicators, which makes dashboards more effective; it encourages the *same* key metrics to be used, based on the same data, for ease of comparison among reporting entities; and it removes an element of guesswork required while already dealing with the many other issues that will confront decision makers during the next pandemic or other public safety crisis

These pre-planned key indicators would be supported by increased funding for data infrastructure, more epidemiologists, and GIS capabilities. While even relatively small entities were sometimes able to create and maintain impressive communication of public health data, the immense technical challenges involved in reporting these data would be better prepared by investments in specialized expertise. In addition, the hundreds of dashboards reporting little useful information besides case numbers shows that many entities could potentially benefit from additional resources. If in-house expertise is not available, consulting with external experts or private partnerships as early as possible may help. Developing and maintaining those relationships ahead of time can be beneficial to move quickly during crisis.

Fourth, entities can prepare more broadly through allocating resources to data visualization needs and creating cultures within and between organizations that are prepared to handle the unexpected challenges which will arise despite the enhanced technical preparation described above. We find empirical evidence showing that the size of entities is strongly correlated with their ability to provide data visualizations to their stakeholders, and that this decision may have been influenced by a variety of local factors such as the presence of other dashboards and demographic characteristics of the local area. In addition to the importance of resources, creating a culture in which securely sharing essential data between governing entities is encouraged may be helpful, as well as creating a culture in which data is available from non-politicized sources, and which can be presented without political agendas.

Fifth, the need for expanding resources for data visualization should be met with lowering the costs and barriers required to provide data visualizations to the public. For instance, the ongoing effort to create public-domain or openly licensed tools which mimic the functionality in more expensive software and can be used by non-specialists should be supported (particularly by philanthropy and/or government), because small entities dealing with public health crises may balk at taking on additional costs, especially during a crisis.[172] Even for paid software options, open-source templates for visualization of public health data may lower the expertise costs required to implement dashboards. For example, the ability of a non-expert to format an Excel or Google Sheet spreadsheet in a certain manner to populate pre-existing

172. *See, e.g.*, *Ten Best Open Source Tableau Alternatives 2022*, RIGOROUS THEMES (Dec. 27, 2021), https://rigorousthemes.com/blog/best-open-source-tableau-alternatives/ [https://perma.cc/DL4P-4N7Y] ("Simply put, there is a lot to gain from using Tableau as a data analyst or business owner. However, just because it is a valuable solution doesn't mean it is ideal for everyone. Mainly because it can be costly for a lot of businesses. Also, users have to be well-versed analysts to take full advantage of its capabilities. Acquiring this level of experience can be equally expensive for business owners, whether by recruiting new staff or training in-house."); *id.* ("[M]any individuals and businesses want alternatives to Tableau and not just any kind. Open source alternatives are cheaper (most are free), and they are more likely to have advanced features that cater to the specific needs of analysts.").

data visualization structures is far less complicated than creating those structures themselves.[173]

To close, one survey respondent's optimistic comment seems salient. When asked about future preparedness, they noted that "[w]e are better prepared already." The process of implementing public communication of COVID-19 data has been a learning experience in which many entities emerged with new knowledge about privacy laws, data organization, community partnerships, the pitfalls of politicizing data, and so on. These lessons learned have certainly strengthened the ability of entities to communicate with their constituents. At the same time, without plans or programs designed to *preserve* these lessons for the future, they may be long forgotten by the time the next inevitable public health crisis emerges.

173. For instance, in Tableau visualization, the ability to create "Sankey" diagrams (flow charts which link two or more categories of data with curved visualizations of the flows between those categories) or curved timelines (visualizations of time data along a line that curves along the screen to capture a longer timespan than would otherwise be possible within the dimensions of a dashboard) are most easily done by finding online templates and then populating them with existing data, rather than performing the host of internal calculations required to create such visualizations from scratch. *See, e.g.*, Ken Flerlage, *Horizontal Sankey Template*, TABLEAU PUBLIC (Sept. 15, 2021), https://public.tableau.com/app/profile/ken.flerlage/viz/Sankey_25/Sankey [https://perma.cc/9GET-E6BZ]; Kevin Flerlage & Ken Flerlage, *Curvy Timelines in Tableau*, FLERLAGETWINS.COM (July 28, 2018), https://www.flerlagetwins.com/2018/07/curvy-timelines.html [https://perma.cc/2Z5N-QJYY].

Appendix: Covid-19 Dashboard Questionnaire

1. Were you involved in the creation, maintenance, management, or evaluation of a dashboard or other data visualization tool that communicated information about COVID-19 to the public?
    a. Yes
        i. What type of entity do you work for (select one)?
            1. Public health department or city, county, state, federal or other government agency
            2. Public or private institute of higher education
            3. Public or private K-12 school or local education agency
            4. News organization
            5. Other _____
        ii. How would you describe your role in communicating information about COVID-19 to the public?
        iii. In your opinion, what were the greatest public benefits from the COVID-19 visualization(s) you created, maintained, managed, or evaluated?
        iv. What were the greatest challenges you faced in your work on communicating information about COVID-19 using data visualization(s)?
        v. What role, if any, did public/private partnerships play in your work on communicating information about COVID-19 using data visualization(s)?
    b. No
        i. Go to question 2.
2. Many different entities such as governments agencies, schools, and news organizations created dashboards or other visualizations related to COVID-19. Please list any specific sources on which you relied for statistical information about COVID-19.
3. What was most effective about the sources you listed?
4. What was least effective about the sources you listed?

5. On a scale of 1 (strongly disagree) to 7 (strongly agree), how would you rate the following statements:
    a. In general, visual communication of information about COVID-19 has been effective in encouraging public behavior that limited the spread of COVID-19.
    b. In general, data used in COVID-19 visualizations has accurately reflected infection rates, hospitalizations, and other key statistics.
    c. Entities across the United States are prepared to communicate information about future public health crises in reliable and effective ways.
6. How has compliance with legal regulations, such as privacy laws or data reporting requirements, affected your efforts to communicate pandemic-related data?
7. Have any of your efforts to communicate accurate, reliable, and complete information about COVID-19 to the public been influenced by political interests?
    a. Yes
        i. In what ways have efforts to communicate accurate, reliable, and complete information about COVID-19 to the public been influenced by political interests?
    b. No
        i. Go to question 8.
8. What are the most important steps that could be taken to prepare for future communication of critical data in a similar crisis to the COVID-19 pandemic?
9. Is there anything else you would like to relate regarding communicating information about COVID-19 through data visualizations?
10. May we contact you for a follow-up interview based on your responses?
    a. No
        i. End of questionnaire
    b. Yes
        i. Please provide an email address which would allow us to follow-up with you directly.

End Page

Thank you for completing this questionnaire. Your responses have been recorded. If you indicated a willingness to participate in a follow-up interview, you will be contacted by the researchers soon.