Charting Public Health: A Taxonomic Study of Visualization Practices in the Public Health Field

Mia Hines and Alvitta Ottley

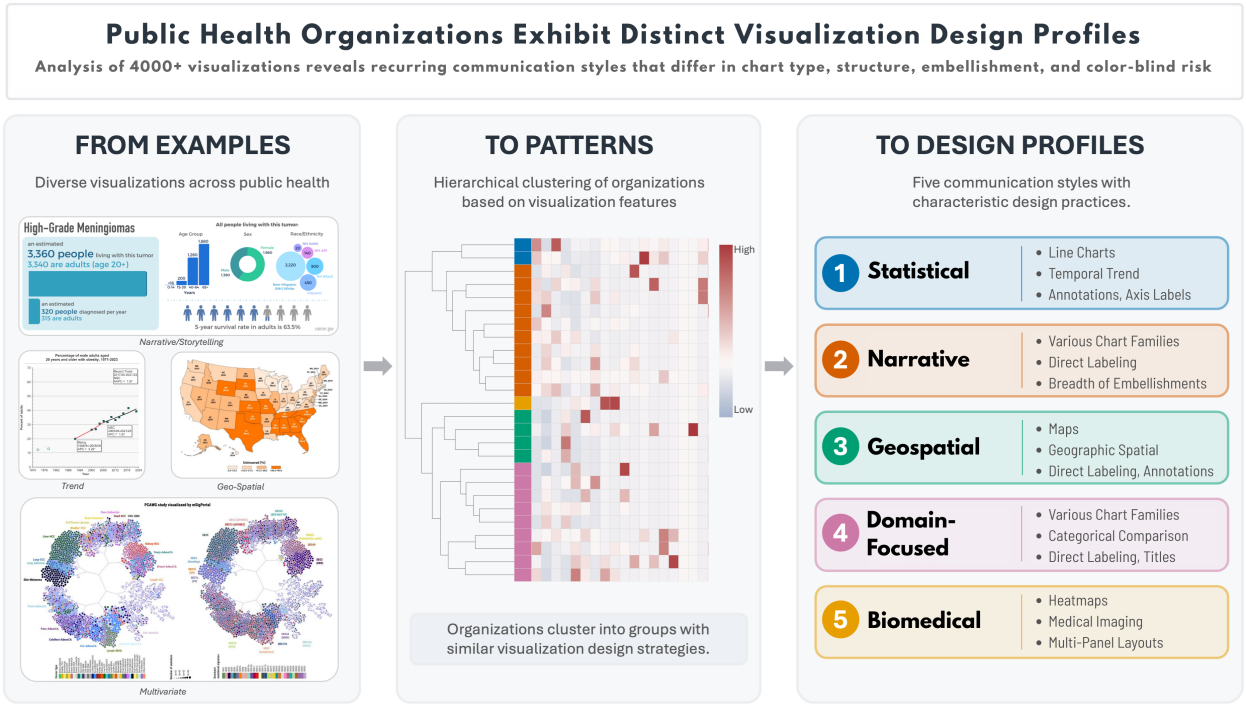


Figure 1: We have clustered organizations across a large corpus of real-world visualizations and uncovered five recurring communication styles that differ in chart forms, structural completeness, annotation practices, and visual embellishments. The resulting profiles reveal systematic differences between evidence-forward reporting, public storytelling, scientific dissemination, and imaging-heavy communication.

Abstract—Public health organizations regularly produce and publish data visualizations to raise awareness of critical issues, influence decision-making processes, and promote overall well-being. However, the design practices shaping these visualizations in real-world settings remain largely unexamined, limiting the research community's ability to evaluate their effectiveness, accessibility, and alignment with communication goals. To address this gap, we construct and analyze a large-scale corpus of over 4,000 real-world data visualizations drawn from more than two dozen websites associated with U.S. and international public health organizations. We evaluate salient design characteristics like chart type, visualization accessibility, use of embellishments like iconography, and design flaws. This work contributes to understanding real-world decisions in designing data visualizations and supports public health officials in improving data visualization-related communications. Visualizations in our finalized corpus and the labeled dataset can be found at https://washuvis.github.io/ChartingPublicHealth/.

Index Terms—Taxonomy, Public Health, Design, Visualization Corpus, Visualization Dataset

1 Introduction

Data visualizations are among the most consequential communication artifacts in public decision-making systems. In domains such as public health, they communicate evidence, shape policy discourse, and influence population behavior. These visualizations appear in a wide range of public-facing materials, including reports, dashboards, infographics, and web content, and play a central role in shaping how health information is understood by policymakers, practitioners, and the general public. For instance, visualizations are consequential for monitoring disease outbreaks, e.g., COVID-19 and HIV/AIDS, and dispatching resources to communities in need [7, 31, 51]. Interactive visualizations are used for electronic health records (EHRs), thereby reducing the cognitive load on medical staff and improving in-patient care [30, 45]. Visualization-focused web applications have been developed to increase shareholder engagement and improve policy-makers' understanding of critical issues [26, 41, 57]. Thus, understanding the design decisions of public health visualizations has direct implications for evaluating comprehension, trust, and action in high-stakes settings [34, 39].

Despite their importance, there is limited empirical understanding

• *Mia Hines is with Washington University in St. Louis E-mail: hines.mia@wustl.edu*
• *Alvitta Ottley is with Washington University in St. Louis E-mail: alvitta@wustl.edu*

Manuscript received xx xxx. 202x; accepted xx xxx. 202x. Date of Publication xx xxx. 202x; date of current version xx xxx. 202x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.202x.xxxxxxx

of how data visualizations are actually designed and deployed in real-world public health settings. Prior work has largely focused on the development of visualization applications [16, 51], literature surveys of specific design factors [37, 40], or has examined practitioner needs through interviews [41]. While these efforts provide valuable insights into what *could* be effective, they offer limited visibility into what is *actually produced at scale*. As a result, fundamental questions remain unanswered: How do organizations communicate data in practice? What design choices dominate real-world deployment? And what tradeoffs or vulnerabilities do these choices introduce?

In this work, we address this gap through a large-scale empirical characterization of real-world visualization practice in the public health domain. We construct and analyze a corpus of more than 4,000 visualizations collected from over two dozen U.S. and international public health organizations. Using a semi-automated pipeline that combines web scraping, computational filtering, and manual validation, we capture a broad, ecologically grounded sample of visualizations as they are encountered by real audiences.

Our analysis reveals a highly concentrated real-world design space. A small number of chart families dominate the corpus, with line charts alone accounting for over 40% of all visualizations and temporal trend communication emerging as the dominant task context. While organizations generally adhere to core structural conventions, we find that important vulnerabilities persist, including inconsistent use of legends and direct labels and widespread color accessibility risks. Notably, nearly one-quarter (23.08%) of visualizations contain potentially risky color pairs under simulated color-vision-deficiency conditions.

Beyond corpus-level trends, we identify distinct *organization-level design profiles*, revealing that institutions exhibit stable, interpretable visualization styles. These styles are likely shaped by subject matter, communication norms, audience expectations, and reporting goals. Although grounded in public health, these findings help expose broader patterns in how organizations operationalize visualization design under real-world constraints and make the following contributions:

- A large-scale, ecologically grounded corpus and annotation framework for studying real-world visualization practice in public health, and provides reusable infrastructure for future research.
- A corpus-level empirical characterization of dominant chart forms, communication tasks, structural conventions, and accessibility vulnerabilities.
- An organization-level analysis revealing stable design profiles that expose how institutional norms shape visualization practice.

2 Related Work

This paper intersects research on real-world visualization corpora, empirical studies of visualization design, and its use in public health and high-stakes communication. We review these areas and highlight key studies that position our work within the expanding research landscape. For a broader overview of corpus-based research in visualization, we recommend the survey by Chen et al. [12].

2.1 Empirical Studies of Real-World Corpora

Prior work in visualization has utilized large collections of real-world visualizations to study design, communication, and automated analysis (e.g., [13, 14, 23, 29, 42]). Borkin et al.'s study of over 2,000 visualizations identified key visual features, such as color, density, and imagery, that enhance memorability [10]. This work had a lasting impact through the release of the *MassVis* dataset, which has become essential across various research areas. It facilitated large-scale studies of visual attention and eye fixation behavior in information visualizations [11], and more recent research on communication, recall, and trust in visualization [4, 34, 39].

Subsequent work has expanded this research by creating large-scale visualization datasets and repositories, such as *VizNet*, which offers a vast collection of visualizations and datasets for machine learning and analysis of visualization structures [21]. *VisImages* enhanced this direction with a detailed, expert-annotated corpus for in-depth structural and semantic analysis of visualizations. [17]. *Beagle* showcased automated extraction and interpretation of web visualizations, facilitating large-scale recovery of visual artifacts beyond curated datasets [6].

A parallel line of work has used real-world visualization collections to study public communication and misinformation. For instance, Zhang et al. mapped COVID-19 crisis visualizations to analyze their use during the public health emergency [58]. Lee et al. explored how visualization practices in online discourse can support unconventional scientific claims by utilizing traditional conventions [28]. Others examined "misleading" visualization practices to study how deception might manifest in real-world artifacts and communicative contexts [32]. These studies show that real-world corpora can reveal structural properties of visualization design and their social and communicative roles.

2.2 Embellishment

Visualization research has explored how design features beyond core data encoding affect interpretation and engagement. A key focus is on *visual embellishments*, i.e., design elements unnecessary for decoding data and often linked to Tufte's concept of "chartjunk" [53]. These include color, direct labeling, iconography, imagery, and explanatory text.

Early perspectives focused on minimalist design with high data-ink ratios to enhance analytical clarity [53]. However, research has shown that embellishment can also play a significant role. For example, Bateman et al. found that embellished charts improved recall without significantly impacting comprehension accuracy [5]. Borkin et al. showed that color, recognizable objects, and high visual density contribute to memorability [10]. Other works have shown that infographic-style visualizations are often perceived as more engaging and aesthetically appealing than plain charts [1], and that design choices impact perceived credibility [49] and clarity [39].

At the same time, prior work highlights important tradeoffs. Certain embellishments can reduce clarity, introduce ambiguity, or negatively affect trust depending on the visual style and communicative context [1, 49]. For example, Song et al. found that comic-style visualizations and hand-drawn font treatments reduced credibility relative to more conventional visual forms [49]. Similarly, Skau et al. found that shape-based embellishments did not improve performance over standard bar charts [48]. Borgo et al. found that visual embellishments improved information retention but increased processing times [8]. Thus, the effects of embellishment are highly context-dependent and often reflect a tension between engagement, memorability, and interpretability.

2.3 Structure & Accessibility

Previous work highlights the need for structural clarity and accessibility in visualizations (e.g., [18, 22, 25]). Essential design elements such as axes, legends, labels, annotations, and color accessibility are vital for accurate interpretation and inclusive communication, particularly in public-facing, high-stakes contexts where simplicity, readability, and precision must be balanced [38]. Moreover, prior work has shown that titles, labels, annotations, and other structural elements influence viewers' attention, perceived informativeness, confidence, and interpretation of visualizations [3, 9, 20].

Accessibility has likewise become central to visualization research. Researchers have developed auditing frameworks such as *Chartability* [18], evaluated accessibility tools and interaction techniques for people with low vision and blindness [24], and expanded the discussion to broader disability communities [56]. Accessibility is crucial in public health, as organizations must adhere to regulations like the Americans with Disabilities Act with the Web Content Accessibility Guidelines (WCAG). Our analysis centers on color vision deficiency (CVD) while supporting future efforts to expand accessibility evaluations.

2.4 Visualization in Public Health Practice

Research on data visualization in public health has mainly focused on applications and systems. Several studies have developed tools for tasks like disease surveillance, intervention targeting, cohort analysis, and public health communication [16, 27, 46, 51, 54]. For example, Sullivan et al. introduced *AIDSVu*, an interactive map for HIV prevention and

Table 1: Public Health organizations within the final corpus.

Organization Name	Acronym	Type	Images Collected	Final Visualizations
Agency for Toxic Substances and Disease Registry	ATSDR	National Agency	923	97
Alzheimer's Association	AA	Nonprofit Organization	1,118	10
American Cancer Society	ACS	Nonprofit Organization	546	29
Centers for Medicare & Medicaid Services	CM&MS	National Agency	360	33
Dana-Farber Cancer Institute	DCI	Nonprofit Organization	1,018	27
Fogarty International Center	FIC	National Agency	1,248	25
Health Resources and Services Administration	HRSA	National Agency	133	8
Indian Health Service	IHS	National Agency	1,901	24
Michael J. Fox Foundation for Parkinson's Research	MJFFPR	Nonprofit Organization	138	2
National Cancer Institute	NCI	National Agency	39,009	2,830
National Eye Institute	NEI	National Agency	1,031	9
National Heart, Lung, and Blood Institute	NHLBI	National Agency	77	5
National Institute of Allergy and Infectious Diseases	NIAID	National Agency	2,486	12
National Institute of Biomedical Imaging and Bioengineering	NIBIB	National Agency	680	22
National Institute of Deafness and Other Communication Disorders	NIDOCD	National Agency	122	44
National Institute of Dental and Craniofacial Research	NIDCR	National Agency	235	3
National Institute of General Medical Sciences	NIGMS	National Agency	1,076	57
National Institute of Health	NIH	National Agency	16,618	718
National Institute of Mental Health	NIMH	National Agency	1,126	82
National Institute on Minority Health and Health Disparities	NIOMHHD	National Agency	211	3
Office of Inspector General	OIG	National Agency	229	56
Pan American Health Organization	PAHO	International Organization	5,967	3
Providence St. Joseph Health	PSJH	Nonprofit Organization	77	3
St. Jude Children's Research Hospital	SJCRH	Nonprofit Organization	9,247	22
Substance Abuse and Mental Health Services Administration	SAMHSA	National Agency	180	143
World Food Programme	WFP	International Organization	3,474	18

care in the U.S. [51, 54]. Cocoros et al. developed a visualization platform for public health surveillance in Massachusetts [16]. Other research has focused on software libraries for creating health-related visualizations [27, 47].

Research on visualization in public health includes literature surveys, scoping reviews, and practical analyses, focusing on decision-making, chart types, dashboard design, electronic health records, and health communication (e.g., [36, 37, 40, 52, 55]). Ofori et al. surveyed 28 articles on high-level design choices for visualization tools in public health [37]. Arleo et al. evaluated COVID-19 dashboard designs and provided expert recommendations [2]. Riddell et al. reviewed 84 papers on network methods and visualizations in public health data [44]. Preim et al. examined visual analytic solutions in public health, highlighting techniques and use cases [43].

Most relevant to this work, Park et al. conducted a qualitative study with 14 public health officials to explore the methods, needs, and challenges of creating visualizations in practice [41]. Their findings highlighted the use of visualizations for communication with health department staff, policymakers, stakeholders, and the public, with a preference for simple, easily interpretable formats. Participants emphasized the effectiveness of infographics for narrative communication and behavior change, as well as maps for displaying geographic distributions of public health issues. However, they also faced challenges such as difficulty accessing appropriate data, inconsistent data sources, and limited staff expertise in visualization design.

In summary, prior research has explored visualization design through experiments and practitioner perspectives, but it *lacks a systematic, large-scale understanding of actual design practices.* There's limited knowledge about how design features co-occur, how choices differ across organizations, and how real-world constraints affect visualization production. This work aims to fill that gap with an empirical analysis of public health visualizations.

3 Method

We collected a large-scale, ecologically grounded corpus and empirically characterized dominant chart forms, communication tasks, structural conventions, and accessibility issues. We gathered **91,414** candidate images and refined them to **4,285 visualizations** through a semi-automated filtering process. We developed a codebook, annotated the corpus using GPT-5 and complementary algorithms, and manually validated random samples after each major processing stage. Figure 2 summarizes the collection and filtering pipeline.

3.1 Corpus Scope and Source Selection

1 We began by compiling an initial candidate set of **57 public health organizations** representing a diverse cross-section of public-facing health communication, including nonprofit organizations, international and multilateral agencies, and U.S. national public health institutions. We selected U.S. and international organizations because of their global influence and their relevance to the literature reviewed. To construct this sampling frame, we conducted targeted Google searches and drew organizations from publicly available directories. We used five directories, including the 2025 Forbes Top Charities and NIH Institutes and Centers; see Appendix A for the complete list of sources and URLs. Organizations were included only if their primary mission focused on health, medicine, or public health.

The resulting candidate set included 18 nonprofit organizations, 9 international and multilateral organizations, and 30 U.S. national agencies. For the purposes of this study, we treated sub-agencies as distinct organizations when they maintained separate public-facing websites and independently curated visual communication materials. For example, the National Institutes of Health (NIH) and the National Institute of Dental and Craniofacial Research (NIDCR) were treated as separate organizations because they operate distinct websites and publish separate visual content, even though NIH is the parent agency of NIDCR. We then applied exclusion criteria to retain only organizations that could be systematically incorporated into the corpus. Organizations were excluded if their websites were inaccessible due to security restrictions or firewalls ($n = 14$), incompatible with our scraping pipeline (e.g., dynamic content, unsupported rendering frameworks, or non-standard image delivery mechanisms; $n = 5$), or yielded no images after automated collection ($n = 6$). After applying these criteria, the set consisted of **32** organizations that were successfully incorporated into the image corpus.

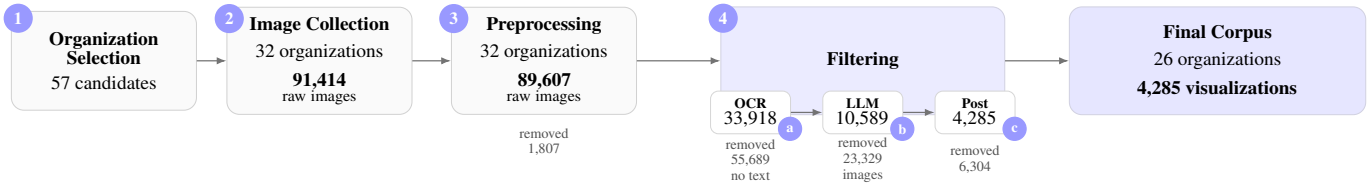


Figure 2: Corpus construction and annotation pipeline. The top-level stages align with the methodological subsections, while nested boxes preserve step-level detail. Starting from 57 candidate public health organizations, we scraped 91,414 raw images and progressively reduced noise through preprocessing and multi-stage filtering, yielding a final corpus of 4,285 visualizations from 26 organizations.

3.2 Web Collection and Image Retrieval Pipeline

2 To support consistent retrieval across heterogeneous website structures, we used a targeted keyword-based search strategy established on terms commonly associated with data communication: *data*, *chart*, *map*, *visualization*, *graph*, *infographic*, *dashboard*, and *plot*. Each keyword was entered into the search function of every organization's website, and all resulting directory index pages were manually collected. For example, searching *data* on the Alzheimer's Association website required manually collecting subsequent index pages (e.g., pages 2, 3, etc.). This process yielded a total of **10,246 candidate URLs**. The scraper then operated iteratively in three stages. First, it visited each collected index page URL. Second, it recursively opened all linked listing pages and content pages accessible from those directories. Third, it extracted all embedded image assets from each visited page.

Data collection was conducted **from January to February 2026** using our self-designed web scraper. Across this process, we retrieved a total of **91,414 candidate images**. In addition to image assets, we collected associated webpage metadata when available, including the page title, description, source URL, page author, and last modified date. These metadata were retained to support future analyses, and summarized in Appendix C.

3.3 File Preparation and Corruption Checks

3 We used the Python Imaging Library (PIL) and `svglib` to standardize image formats and remove unusable files to support downstream computational processing. SVG and WEBP file formats were converted to PNG to ensure compatibility across subsequent analysis steps. We additionally identified and removed corrupted image files that could not be successfully opened or processed. After this stage, 2,522 files were converted and **1,807 images** were removed due to corruption or format issues, yielding **89,607 images** for subsequent analysis.

3.4 Visualization Identification and Multi-Stage Filtering

4 Because public-facing organizational websites contain a substantial number of non-visualization images, the raw image collection included significant noise. The majority of retrieved images were not data visualizations, which is consistent with prior work on web scraping public-sector and news websites [10, 33]. The collected noise in our study was software screenshots, logos, promotional flyers, anatomical diagrams, document pages, organizational charts, and photographs. To construct a high-quality visualization corpus, we implemented a four-stage filtering pipeline that progressively reduces noise while preserving candidate visualizations.

4a **OCR-Based Text Presence Filtering.** We applied an optical character recognition (OCR) filter, assuming that data visualizations contain some textual elements, such as axis labels, legends, titles, tick labels, or direct annotations. This stage was designed primarily to remove photographs and other non-informational imagery, which we anticipated to be a major source of the collected noise. We used `Pytesseract`, a Python wrapper for Google's Tesseract OCR engine, to detect text within each image. Images for which no text was detected were removed from the candidate corpus. Additional removals occurred 1) if the image structure wasn't decipherable by OCR, or 2) if the OCR text extraction was empty after applying text rules to remove special characters and spaces. This step removed **55,689 images**, reducing the corpus to **33,918 candidate images**. All extracted OCR text was retained for later use in taxonomy development and metadata enrichment.

Validation An author manually reviewed a random sample of **406** images, achieving **84%** agreement with the automated decision. Disagreements typically occurred because images exhibited bad resolution, poor contrast, or were too small, which would've removed in other stages of the pipeline.

4b **LLM-Assisted Visualization Identification.** We used OpenAI's `gpt-5-mini` model to classify each candidate image using a structured prompt that included the following outputs: whether the image was a visualization, chart type labels, text presence, noise category, structural features (e.g., axes, marks, panels), and a confidence score ranging from 0 to 1. To maximize precision, we retained only images that satisfied all of the following criteria: (1) labeled as a visualization, (2) assigned at least one valid chart type, (3) noise category labeled as `none` or `other`, and (4) model confidence greater than **0.85**. We chose a strict threshold of 85% to have greater confidence that noise was removed from our corpus, but we note that the trade-off is the potential exclusion of valid visualizations. This step removed **23,329 images**, yielding **10,589 candidate visualizations**.

Validation An author manually validated a random sample of **100 images**, with agreement the image contains a data visualization exceeding **97%**. Labels generated during this stage were retained for later design-space analysis. Disagreement typically occurred with hierarchical organizational diagrams, which were intended to be classified as noise but were assigned as a visualization from the LLM.

4c **Post-Processing and Deduplication.** Finally, we applied a post-processing stage to remove residual artifacts that persisted through earlier filters. Images were excluded if they were classified as (1) too small or (2) a duplicate. 1) We considered an image too small if both width and height were less than or equal to **400 pixels**, which, depending on the device, is comparable to the size of a postage stamp. 2) We used `PIL` and the `imagehash` packages to identify exact duplicates using perceptual hashing and retained a single canonical copy for each unique image. 24 images were manually removed because the related organizations met the exclusion criteria described in Section 3.1; the image artifacts were collected before the organizations were observed to meet the exclusion criteria. This final stage removed **6,304 images**, resulting in a final corpus of **4,285 visualizations**. Another **6 organizations** were ultimately removed because their image collections became empty after the reduction process. The final corpus, therefore, consists of **26 public health organizations** and **4,285 visualizations**. Table 1 summarizes the organization-level counts from initial retrieval through final corpus inclusion.

3.5 Codebook Development

We used an iterative codebook development process. Instead of directly applying labels from prior work in visualization taxonomy, the authors conducted an open-coding pass to ensure appropriate capture of design elements in this specialized collection of visualizations. One author conducted an initial open-coding pass over the sampled visualizations, documenting all meaningful annotation attributes observed in each artifact. The sample was constructed by randomly selecting 5 visualizations (if applicable) from each organization in the corpus, resulting

in a total of 131 visualizations reviewed. These attributes spanned multiple dimensions of visualization design, including content (e.g., demographic focus such as gender or race), visual embellishments (e.g., iconography, photographs, decorative elements), and structural features (e.g., chart type, axis labels, legends, annotations). This process initially yielded 30 variables

The full author team then reviewed and iteratively refined the initial codes into a consolidated codebook. During this process, redundant, overly specific, or semantically overlapping attributes were merged into broader coding dimensions. For example, rotated labels, font size, and color scale tags were removed because they were overly specific or unlikely to be consistently captured across the corpus. Rather than annotating individual data types (e.g., categorical, time series, quantitative), we developed a broader *Data Context* taxonomy to characterize the visualization's primary communicative task.

The final annotation schema comprised 22 variables organized into six categories: *chart family*, *data context*, *structural features*, *visual embellishments*, *color*, and *text*. Table 3 summarizes the codebook and details the variables in the Appendix B. The resulting codebook was used to annotate the full corpus and support the subsequent design-space analyses.

3.6 LLM-Assisted Corpus Annotation and Validation

Following codebook development, we operationalized the final annotation schema as two structured prompting frameworks and applied it to the full corpus using OpenAI's `gpt-5-mini` model. The prompts used a controlled vocabulary derived directly from the finalized codebook. The first prompt classified each visualization with *visual embellishments* and *color* features. *Visual embellishments* features used a binary feature schema (e.g., presence of attributes such as iconography). The *color* associated attribute within this prompt was the extraction of the HEX values from the visualization. The second prompt classified each visualization according to *chart family*, *data context*, *structural features*, and a confidence score ranging from 0 to 1. To account for hybrid visual forms, the model was instructed to return both a *primary* and an optional *secondary* chart type. Likewise, the model selected *primary* and *secondary* options for *data context*. To ensure consistency, *structural features* and *data context* attributes were accompanied by explicit natural-language definitions in the prompt. Like visual embellishments, *structural features* attributes follow a binary feature schema (e.g., denoting the presence of direct labels, legends, and grid lines).

The model was required to return all outputs as structured JSON. Some fields were excluded from subsequent analyses. Secondary chart-family labels were omitted because GPT assigned over 50% of visualizations to the generic "Other" category, preventing meaningful analysis of hybrid visual forms. For consistency, secondary data-context labels were also excluded. The confidence score from the second annotation prompt was removed from forthcoming analyses. We adopted a conservative confidence threshold of 87% to guide quality assurance. We emphasize that this threshold was used only to prioritize manual review and not as a decision boundary for the final dataset. Because over 75% of visualizations exceeded this threshold, confidence scores were not included in the following analyses. The final annotated corpus contains structured labels for chart family, data context, structural features, visual embellishments, and color values for all **4,285** visualizations. *The corpus, codebook, GPT prompts, analysis scripts, and archived URLs are available* at `https://washuvis.github.io/ChartingPublicHealth/`.

Validation An author manually reviewed a random sample of **100** visualizations in the final corpus. Agreement was assessed for features relating to *chart family*, *data context*, *structural features*, *visual embellishments*, and *color*. Across the validation sample, agreement was **98%** for primary chart type, **98%** for primary data context, and **98%** for HEX colors. In consideration of all structural features and visual embellishments attributes, **95%** average agreement for structural features, and **90.6%** average agreement for visual embellishments. Regarding visual embellishments, disagreements typically arose from the over-application of a particular feature from the GPT model. For instance, visualizations that listed the data sources were sometimes denoted as having text annotations.

3.7 Derived Measures and Communication Risk Indicator

We derived a set of quantitative measures designed to capture higher-level properties of visualization communication and accessibility, such as the concentration of embedded text and potential accessibility risks associated with color palette selection.

3.7.1 Text Density.

To quantify the amount of written information embedded in each visualization, we computed a normalized text density measure based on optical character recognition (OCR). For each image, text was first extracted using Tesseract OCR (as described in Section 3.4). To reduce noise in our text density computation, symbols, whitespaces, punctuation, and numbers were removed from the extracted text. We created tokens by aggregating all the characters left in the text. We then normalized text quantity by image area (width and height) to account for differences in visualization size:

$$D_{\text{text}} = \frac{N_{\text{char}}}{W \times H}$$

where N_{char} is the total number of retained text characters, and W and H denote the image width and height in pixels, respectively. This normalization yields a unitless measure of text density that enables comparison across visualizations of varying dimensions, where larger values indicate a greater concentration of embedded textual information relative to image size.

3.7.2 Color Blindness Risk

To assess potential accessibility concerns, we derived a binary *color blindness risk* indicator for each visualization based on the extracted color palette. For each unique set of detected hex color values, we evaluated whether the palette remained perceptually distinguishable under normal vision and common forms of color vision deficiency (CVD), including protanomaly, deuteranomaly, and tritanomaly.

Each palette was first converted from hexadecimal color codes into normalized sRGB values. We then simulated three common CVD conditions using a perceptual color space transformation implemented with `colorspacious`. For each condition, we computed pairwise Euclidean distances between all color pairs in CIELAB space, using ΔE_{00}, which provides a perceptually grounded estimate of color distinguishability by accounting for differences in lightness, chroma, and hue.

For each visualization, we determined palette separability using the minimum pairwise perceptual distance between all colors:

$$d_{\min} = \min_{i,j,c} \Delta E_{00}^{(c)}(i,j)$$

where i and j index color pairs and c indexes normal vision and the three simulated CVD conditions.

A palette was classified as *risky* if at least one pair of colors exhibited limited perceptual separation under any viewing condition. Following common practical interpretation thresholds, we considered palettes with $\Delta E_{00} < 5$ to indicate elevated accessibility risk, and those with $\Delta E_{00} < 2$ to indicate near-indistinguishable color pairs. See Figure 5 in Appendix D for more details.

For the binary indicator used in our primary analysis, a visualization was marked as risky if:

$$R = \begin{cases} 1 & \text{if } d_{\min} < 5 \\ 0 & \text{otherwise} \end{cases}$$

4 Corpus-Level Results

In this section, we characterize the visualization design space through analyses of chart families, data contexts, structural features, visual styling, and derived communication measures, including text density, color usage, and accessibility risk.

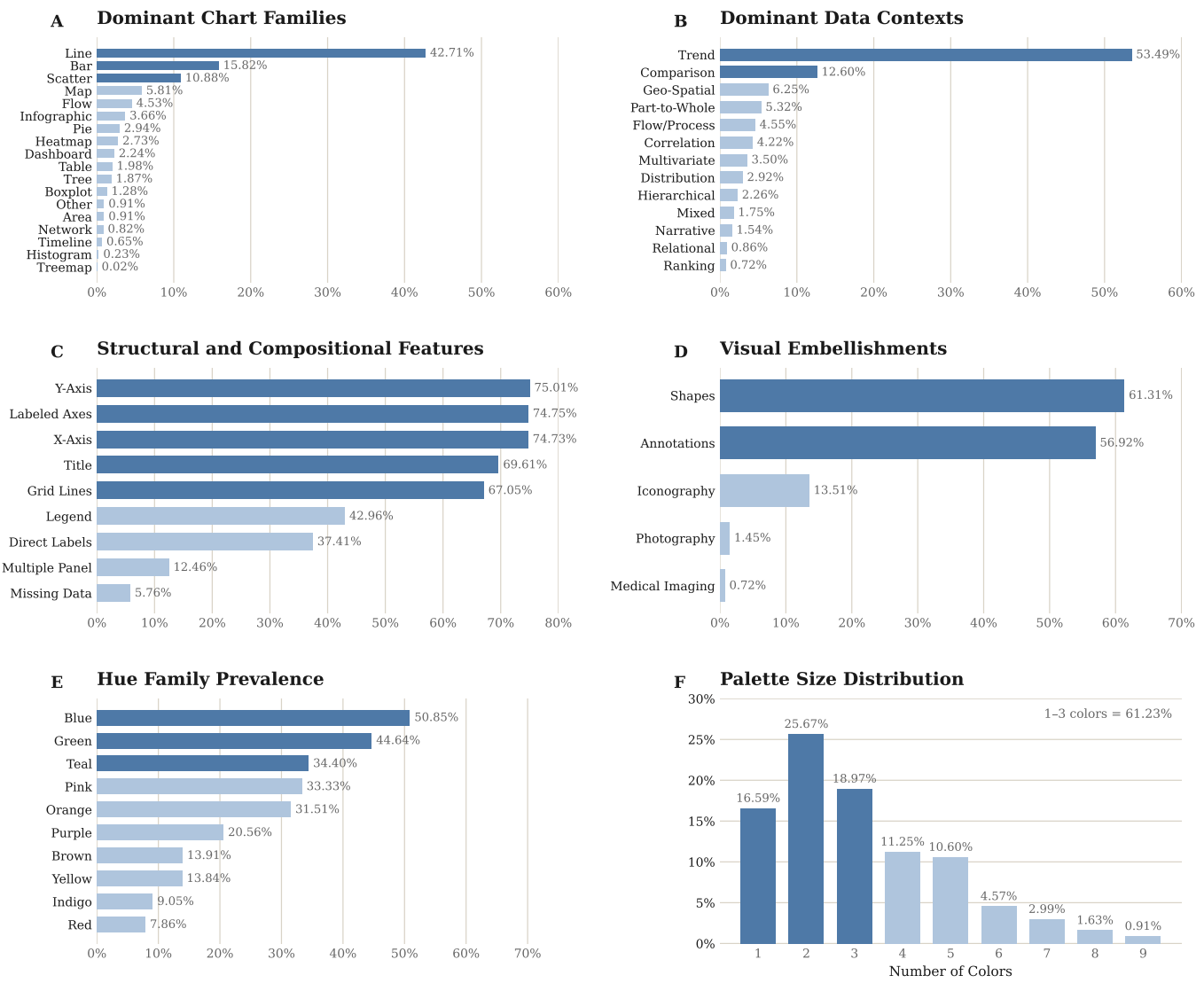


Figure 3: A summary of the analysis findings. Line charts and temporal trends are most prevalent in the corpus, suggesting a strong emphasis on communicating change over time. Chart elements such as axes, labels, and titles are present in most visualizations, while direct labeling is less common. Text annotations and shape/line-based embellishments appear in most charts, indicating that "embellishments" are common in practice.

4.1 A Small Set of Chart Types

As shown in Figure 3A, line charts account for 42.71% of the corpus, substantially exceeding all other chart types. Bar charts (15.82%) and scatter plots (10.88%) are the next most prevalent forms. Together, these three chart families comprise nearly 70% of the entire corpus. The remaining design space is comparatively sparse. Maps account for 5.81% of visualizations. Infographics (3.66%), dashboards (2.24%), tables (1.98%), and more structurally specialized representations such as treemaps (0.02%) appear infrequently. This long-tailed distribution suggests that public health communication relies heavily on a small, stable set of familiar chart forms. *We observe limited chart diversity, with organizations overwhelmingly favoring conventional charts over more specialized or exploratory designs.*

4.2 Temporal Trend Communication Dominates

We can observe from Figure 3B that the dominant chart families align closely with the primary data contexts observed in the corpus. Temporal trend communication is by far the most common analytical context, accounting for 53.49% of all visualizations. This prevalence substantially exceeds the categorical comparison (12.6%) and the geographic context (6.25%). Other contexts, including part-to-whole composition (5.32%), flow processes (4.55%), and correlation relationships (4.22%), occur substantially less frequently. More complex contexts, such as hierarchical (2.26%) and mixed multi-context visualizations (1.75%), are rare. *The corpus is strongly oriented toward communicating change over time, consistent with the prevalence of line charts.*

4.3 Obscurity in Structural Conventions

Structural elements associated with readability and interpretability appear frequently but not universally. Figure 3C shows that titles are present in 69.61% of visualizations, and grid lines appear in 67.05%. Axes and axis labels are present in approximately three-quarters of the corpus, with Y-axis presence at 75.01%, X-axis presence at 74.73%, and labeled axes at 74.75%. However, other critical interpretive structures appear considerably less often. Legends are present in only 42.96% of visualizations, and direct labels in 37.41%. In visualization design, they often occur in opposition of each other, e.g., legend or direct labeling, and to have both present may be seen as redundant. This is

demonstrated by 69% of visualizations having direct labeling or legends, and approximately 12% exhibiting both features. Multiple-panel designs remain relatively uncommon (12.46%), and explicit denotation of missing data appears in only 5.76% of cases. *While public health organizations generally adhere to core structural conventions, the limited use of legends and direct labels may create opportunities for interpretive ambiguity in public-facing communication.*

4.4 Embellishments Are Functional

Figure 3D shows that lightweight embellishments are common, particularly unique point shapes and line markings (61.31%) and text annotations (56.92%), suggesting that organizations frequently use minimal stylistic variation to support differentiation and emphasis. More expressive embellishments, however, remain uncommon. For example, iconography appears in only 13.51% of visualizations.

Using the normalized text density measure defined in Section 3.7.1, we examined text usage in the corpus. Text density varied substantially across the corpus, suggesting notable differences in communication style across organizations. The bottom 25% of visualizations contained very little embedded text ($\leq 1.33 \times 10^{-4}$), indicating highly minimalist designs. By contrast, only the top 25% of visualizations exceeded a text density of 1.70×10^{-3}, suggesting that text-rich explanatory designs were comparatively uncommon. Most visualizations balanced graphical encoding with modest amounts of supporting text, *exhibiting a restrained but consistent use of visual styling features.*

4.5 Color Usage Reflects Strong Palette Conventions

Figure 3E suggests that color usage across the corpus reveals strong and highly consistent palette conventions. Blue is by far the most prevalent hue, appearing in 50.85% of all visualizations, followed by green (44.64%) and teal (34.4%). Secondary colors such as pink (33.33%), orange (31.51%), and purple (20.56%) also appear frequently, while warm and high-salience hues such as red (7.86%) and yellow (13.84%) are comparatively rare. In addition to hue selection, palette size is similarly constrained. Our computed palette size only considers hue color, e.g., blue, green, or teal, and not grayscale-associated colors, e.g., gray, white, or black. We can observe from Figure 3F that the most common palette sizes are two-color (25.67%) and three-color (18.97%) designs, followed by single-color visualizations (16.59%). Together, visualizations using one to three colors account for over 61% of the corpus. Larger palettes are substantially less common, with palettes of more than 5 colors occurring infrequently. In sum, the visualizations in the corpus rely heavily on cool-toned palettes, particularly blue and green hues, which may reflect both institutional branding conventions or the perceived neutrality, professionalism, and trustworthiness often associated with these colors [19, 35, 50]. *Consistent with the previous analyses, the limited chromatic complexity indicates a preference for simplicity in public health visualization design.*

4.6 Potential Color-Blind Accessibility Risks

Because public health visualizations are often designed for broad public audiences, accessibility is a critical dimension of communication effectiveness. In particular, inadequate color separation can substantially reduce interpretability for viewers with common forms of color vision deficiency, potentially obscuring trends, categories, or comparisons that are central to decision-making. To assess how frequently such risks appear in practice, we evaluated the corpus using the *color blindness risk* measure described in Section 3.7.2.

We found that color blind accessibility concerns persist in a nontrivial portion of the corpus. Using a practical perceptual distinguishability threshold of $\Delta E_{00} < 5$, *23.08%* of visualizations contained at least one potentially ambiguous color pair under normal vision or simulated color vision deficiency conditions. Even under a stricter threshold of $\Delta E_{00} < 2$, *6.18%* of visualizations contained near-indistinguishable color pairs. *Together, these findings suggest that color-based ambiguity remains a meaningful accessibility concern in real-world public health visualizations, particularly for viewers with color vision deficiencies.*

5 Organization-Level Design Profiles

To examine whether organizations exhibit distinct visualization practices, we clustered each organization's aggregated visualization profile. For each organization, we computed the proportion of visualizations exhibiting each *chart family* (CF), *data context* (DC), *structural feature* (SF), and *visual embellishment* (VE). *Color* and *text* variables were excluded because they represent derived measures rather than GPT-assigned annotations. This produced an organization-by-feature matrix describing each organization's visualization practices.

All features were standardized (z-scores) prior to clustering. We then applied hierarchical agglomerative clustering with Ward linkage to identify *organizational visualization styles*. A five-cluster solution was selected based on dendrogram structure and interpretability, as lower cut points merged distinct styles while higher cut points primarily over-segmented existing ones.

5.1 Organizations Exhibit Distinct Design Profiles

Figure 4 shows the resulting clustered heatmap and a PCA projection of the same standardized feature space. Together, these views reveal **five distinct organization-level design profiles**, suggesting that institutions exhibit stable and interpretable visualization styles. Cluster labels are STATISTICAL, NARRATIVE, GEOSPATIAL, DOMAIN-FOCUSED, and BIOMEDICAL. The labels are descriptive and based on a visual inspection of representative charts. The underlying clusters were formed using low-level visual features extracted from the figures, while the labels (e.g., Statistical and Narrative) summarize the dominant visual characteristics observed within each cluster. In contrast, the PCA projection provides a complementary view of similarity at the organizational level, which does not necessarily align perfectly with the visually derived clusters. Figure 6 showcases example visualizations for each design profile in Appendix E.

STATISTICAL The first cluster comprises of **2,887 visualizations** across **2 organizations**, NCI and NIGMS. Visualizations in this profile exhibit features related to statistical analysis, e.g., regression lines, confidence intervals, p-values, and statistical significance annotations. In consideration of chart family, STATISTICAL has a strong preference for *line charts* at 55.6%, with scatter plots being the second dominant choice at 15.2%. The salient data context is *temporal trends* at 63.9%, which is a rate similar to the preferred chart forms. The design profile values *high structural completeness*, with the labeled axes (85.9%), x-axis (85.2%), y-axis (85.1%), grid lines (78.7%), and titles (77.1%) all above 75%. *Pictorial-related visual embellishments are rare.* Instead, their preferred embellishment features are unique point shapes and marked lines at 71.3% and text annotations at 65.4%. The prevalence of the other visual embellishment features is below 9%. *Overall, this design profile emphasizes standardized statistical reporting and an evidence-based communication style.*

NARRATIVE The second cluster comprises **1,155 visualizations** and **10 organizations**, including NIH, SJCRH, and ACS. Visualizations in this profile demonstrate design decisions related to mass communication and narrative. This is shown in its *wide range of chart families* (e.g., flow diagrams, tree diagrams, infographics, and dashboards) and the use of explanatory elements (e.g., labels, legends, and text annotations). Its most prevalent chart family is a *bar chart* at 37.3%, followed by line charts at 17.3%. In consideration of data contexts, *temporal trends* and categorical comparison are favored at 35.5% and 21%, respectively. A *variety of structural features are exhibited*, with direct labeling (60%), the presence of y and x axes (≈57%), axis labels (54.7%), title (49.9%), grid lines (44.6%), and legend (39.3%) approximately above or equal to 40%. Like structural features, the *use of visual embellishments is broad*, with unique point shapes and marked lines at 39.9%, text annotations at 34.2%, and iconography at 20.3%. *The design profile utilizes a broad range of chart forms, structural features, and visual embellishments with the purpose of using the most fitting combinations to support the high-level narrative from the data. This further suggests that the* NARRATIVE *communications are produced for a myriad audience, e.g., stakeholders, staff, and the general public.*

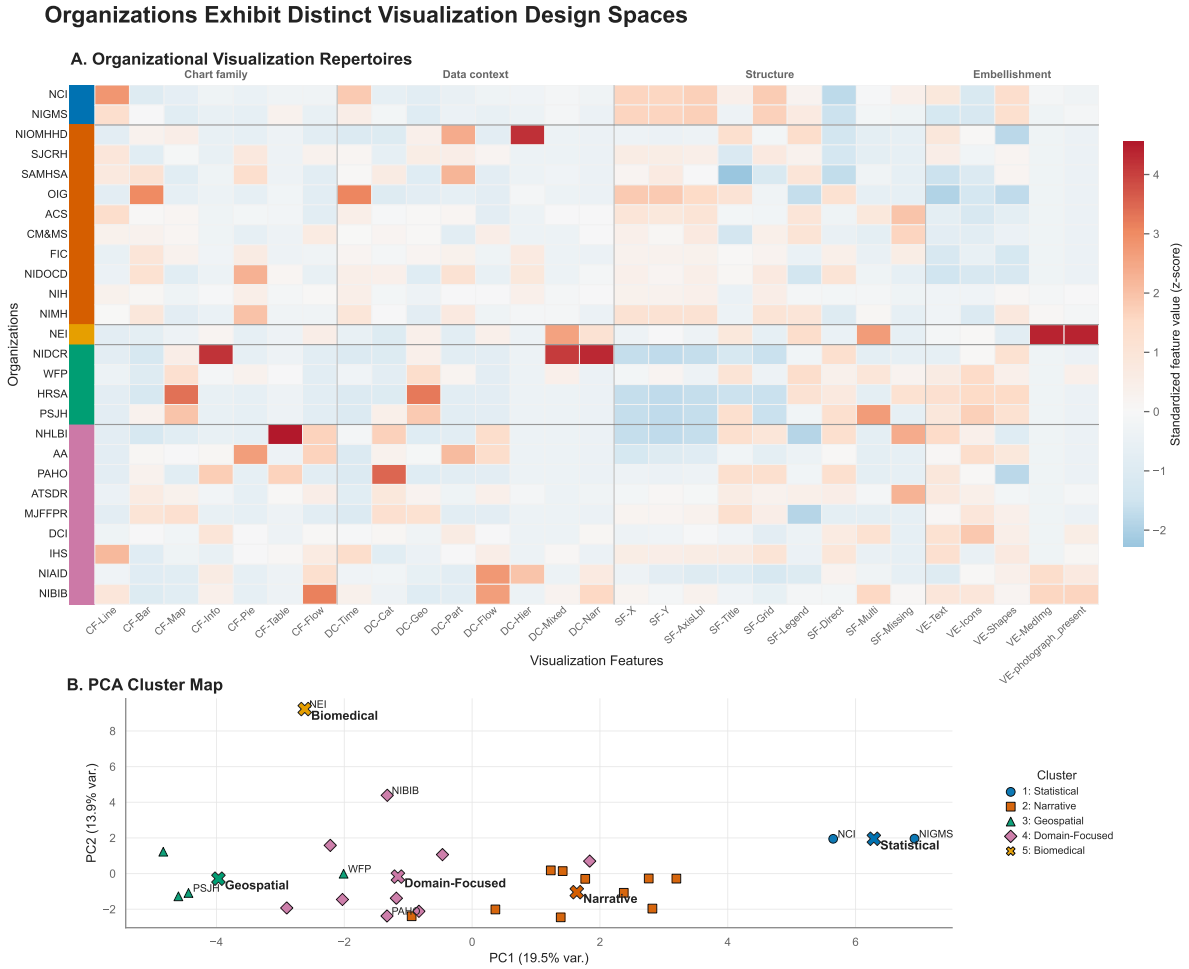


Figure 4: Organizational visualization design space revealed through hierarchical clustering. (A) Clustered heatmap of standardized feature prevalence across chart family (CF), data context (DC), structural features (SF), visual embellishments (VE), and text density (TD). (B) PCA projection of the same feature space, showing separation among the five identified clusters.

GEOSPATIAL The third cluster comprises **32 visualizations** and consists of **4 organizations**: WFP, PSJH, NIDCR, and HRSA. Its visualizations predominantly relate to geospatial data, resulting in various map types (e.g., regional, choropleth, weather tracking, and facility). As mentioned, *maps* and *geographic spatial* are the salient chart family and data context *geographic spatial* features at 62.5% and 65.6%, respectively. Its predominant structural features are *direct labeling* at 84.4% and *titles* at 71.9%. The profile utilizes *visual embellishments consistently*: text annotation prevalence is at 62.5% with unique point shapes and marked lines and iconography both at 59.4%. *The design profile predominantly consists of map visualizations and utilizes direct labeling and visual embellishments, e.g., text annotations and iconography, for clarity in their communication practices.*

DOMAIN-FOCUSED The cluster comprises **202 visualizations** and consists of **9 organizations**: NIBIB, IHS, and PAHO. Under visual inspection, we observed that visualizations in this profile represent domain-specific phenomena (e.g., MRI images, spectra, movement tracking, microscopy, and pathology), suggesting that the intended audience is limited to the organization and experts related to their specific health topic. The profile employs *a broad range of chart forms*, with *bar charts* (27.7%) and *line charts* (12.4%) as most prevalent. Favored data contexts are *categorical comparison and geographic spatial* at 31.2% and 19.3%, respectively. The profile exhibits a wide range of structural features, but has a strong preference for *direct labeling* at 81.2% and the *presence of titles* at 74.3%. Like structural features, a variety of visual embellishments are used but greatly favor *text annotations* at 64.9%. Like NARRATIVE, the design profile utilizes a variety of chart forms, structural features, and embellishments, but implores explanatory design elements (e.g., text annotations and direct labeling) at an increased rate. *The profile characteristics reflect scientific communication that balances standardized presentation with attention-guiding elements.*

Finally, BIOMEDICAL consists of **9 visualizations** and **1 organization**: the National Eye Institute (NEI), indicating a highly distinctive visualization profile. It is characterized by frequent use of *heatmaps* (33.3%) and the data context of *geographic spatial* (33.3%). Salient structural features are *titles (89%)*, *multi-panel layouts (67%)*, and *legends (67%)*. The profile's favored visual embellishment is *text annotations* (44.4%). Its defining characteristic is the use of *medical imaging (22%)* distinguishing it from all other design profiles and reflecting its specialized communication needs.

Importantly, these profiles are not determined solely by subject matter. For example, cancer-related organizations span multiple clusters, indicating that organizations develop distinct ways of visualizing that reflect communication norms, intended audiences, and reporting goals in addition to the underlying content.

5.2 Diverse Color Practices

Color features were not included in our PCA groupings because they are secondary measures; however, we still evaluated their consistency and contrast across all organizations in the corpus. We evaluated the presence of each hue category (e.g., blue, green, and indigo) across all visualizations of an organization. 23% of organizations use a particular color across all visualizations, indicating a consistent color palette or strong brand identity, such as blue for the Health Resources and

Services Administration (HRSA) and orange for the Michael J. Fox Foundation for Parkinson's Research (MJFFPR). Blue is the dominant color choice for 73% of organizations. The size of the color palette in a visualization typically varies across an organization's corpus, with only a few demonstrating strong consistency in palette size. The Office of Inspector General (OIG) exhibits the highest consistency in color palette size, with 91% of visualizations using a single hue. Grayscale visualizations are rarely produced across organizations, and many don't have a single one. *Organizations' consistency in color selection shows that color may be used to enforce a visual identity.*

6 Discussion

Many of our findings are unsurprising. Public health organizations favor familiar chart types, communicate primarily through temporal trends, generally follow established structural conventions, and still exhibit accessibility challenges. Together, however, they suggest something more interesting: *the forces shaping visualization practice may operate less at the level of individual charts and more at the level of organizations.*

Looking at a single visualization naturally leads us to ask why a designer chose a line chart, omitted direct labels, or selected a particular color palette. Implicitly, we treat every visualization as an independent design decision. Looking across more than 4,000 visualizations tells a different story. The observed consistency in chart types, annotations, color usage, and organization-level design profiles suggests that many of these decisions are not made from scratch. Instead, they emerge from recurring organizational processes. This shift in perspective changes the unit of explanation. Rather than asking why this chart looks the way it does, it prompts us to examine why organizations repeatedly produce such charts. The answers are unlikely to be limited to visualization principles but also extend to communication goals, reporting workflows, templates, branding, governance, and institutional memory.

6.1 Organizations Develop Stable Ways of Visualizing

Viewed through this lens, our findings become less about individual design choices and more about stable organizational practices. Organizations appear to develop recurring visualization practices that persist across reports, topics, and authors. The stability of design profiles suggests that organizations repeatedly solve similar communication problems using an established visual vocabulary. Likewise, recurring annotation practices and restrained color palettes likely reflect communication habits that have evolved over years of organizational practice rather than decisions made anew for every visualization. Branding requirements, inherited templates, software defaults, production timelines, and review processes all shape the final visualization. Similar questions can be asked of chart diversity, labeling conventions, and other design recommendations.

This perspective resonates with recent work on visualization style guides [15,38]. Ottley et al. analyzed over 2,000 guidelines from across 52 organizations and found that guidelines are shaped by values such as clarity, efficiency, comprehensibility, and simplicity [38]. Choi et al. evaluated 226 guidelines and found that guidelines are commonly focused on problem-solving and enhancement [15]. These works further elaborate that practice can be seen as a product of an organization's values and culture. Furthermore, style guides explicitly document how organizations believe visualizations should be created, while the present work examines what organizations actually publish. Together, they suggest that organizations develop stable ways of visualizing that are both codified and enacted through everyday practice. Understanding how these practices emerge, evolve, and sometimes diverge from published guidance remains an important direction for future work.

6.2 Closing the Loop Between Research and Practice

Visualization research has traditionally asked how to design better charts. Our findings suggest an equally important question: How do better charts become standard practice? Understanding that process may be a grand challenge for empirical visualization research.

Evaluating departures from best practices is an important investigation in visualization research. Accessibility issues, for example, can be interpreted as poor design decisions. At the organizational scale, they instead point to a different question: *Why do accessibility recommendations fail to become routine practice?* The answer is likely more complex than individual designer expertise. Joyner et al. conducted an in-the-wild analysis of visualization accessibility and interviewed over 100 practitioners [22]. Their findings identified three levels of uncertainty in the designer's evaluation of a visualization's accessibility: personal, operational, and organizational. The work further illuminates the complexity of visualization practitioners' decisions beyond skill level, chosen topic, or specific data task.

7 Data Availability and Ethical Considerations

The corpus comprises visualizations collected from 26 public health organizations, the majority of which are U.S. government agencies whose visualizations are in the public domain. Data collection was limited to publicly available web pages requiring no authentication, and organizations with detectable security restrictions (e.g., Cloudflare) were excluded from the corpus. Accordingly, we will release the complete set of public-domain visualizations, along with the derived annotations, metadata, codebook, analysis scripts, and archived source URLs, creating a reusable resource for visualization research. Visualizations from nonprofit and international organizations will not be redistributed because they may remain the intellectual property of their respective organizations; in these cases, we provide metadata, features, and archived source URLs to facilitate retrieval from the original sources where permitted. Researchers using the corpus for downstream applications, including computational model development, should ensure that their use complies with applicable copyright, licensing, and website terms. This study did not involve human participants and therefore did not require Institutional Review Board (IRB) approval.

8 Limitations and Future Work

Our work has several limitations. First, while the corpus is large and ecologically grounded, it is restricted primarily to public-facing visualizations from U.S. public health organizations and does not capture many international, non-English, internal, or interactive visualizations. Second, the codebook was initially developed by a single visualization expert before iterative refinement, introducing the possibility of subjective bias in the coding framework. Third, although our automated extraction and annotation pipeline was manually validated, it may still contain classification errors. Our conservative filtering criteria may have excluded some valid visualizations, while automated chart classification assigned a single primary chart family and may not fully represent hybrid or multi-view visualizations. Finally, because no manual review was performed on every collected image, a small number of non-visualization images may remain in the corpus.

Future work should extend this approach beyond public health to additional domains and examine how organizational visualization practices evolve over time, particularly during major public health events or policy shifts. An equally important direction is to connect corpus-level patterns with audience studies, to examine how visualization practices influence comprehension, trust, and decision-making.

9 Conclusion

We presented a large-scale empirical characterization of real-world visualization practice through the lens of public health communication. By constructing and analyzing a corpus of more than 4,000 visualizations collected from over two dozen U.S. and international organizations, we revealed a design space that is highly concentrated, structurally uneven, and institutionally shaped. Our findings show that a small set of familiar chart forms dominates public-facing communication, important accessibility and interpretability risks remain widespread, and organizations exhibit stable design profiles shaped by communication norms and reporting goals. Beyond the domain of public health, this work highlights the value of corpus-scale analysis as a methodological approach for visualization research.

Acknowledgments

This work is supported in part by the National Science Foundation under Grants Nos. 2142977 and 2330245, which support the Engineering Research Center for Carbon Utilization Redesign through Biomanufacturing-Empowered Decarbonization (CURB). M. Hines was supported by the following fellowship programs in sequential order: Clare Boothe Luce and NRT-AI: AI Advancements and Convergence in Computational, Environmental, and Social Sciences (AI-ACCESS) (No. 2244165).

References

[1] T. Andry, C. Hurter, F. Lambotte, P. Fastrez, and A. Telea. Interpreting the effect of embellishment on chart visualizations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–15. ACM, NY, 2021. doi: 10.1145/3411764.3445739 2

[2] A. Arleo, R. Borgo, J. Kohlhammer, R. A. Ruddle, H. Scharlach, and X. Yuan. Reflections on the use of dashboards in the covid-19 pandemic. *IEEE Computer Graphics and Applications*, 45(2):135–142, 2025. doi: 10.1109/MCG.2025.3538257 3

[3] A. Arunkumar, L. Padilla, G.-Y. Bae, and C. Bryan. Image or Information? Examining the Nature and Impact of Visualization Perceptual Classification. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):1030–1040, 2024. doi: 10.1109/TVCG.2023.3326919 2

[4] A. Arunkumar, L. Padilla, and C. Bryan. Modeling and Measuring the Chart Communication Recall Process. In *Computer Graphics Forum*, vol. 44, p. e70099. Wiley Online Library, 2025. doi: 10.1111/cgf.70099 2

[5] S. Bateman, R. L. Mandryk, C. Gutwin, A. Genest, D. McDine, and C. Brooks. Useful junk? The effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2573–2582. ACM, NY, 2010. doi: 10.1145/1753326.1753716 2

[6] L. Battle, P. Duan, Z. Miranda, D. Mukusheva, R. Chang, and M. Stonebraker. Beagle: Automated Extraction and Interpretation of Visualizations from the Web. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–8. ACM, NY, 2018. doi: 10.1145/3173574.3174168 2

[7] L. Baxter, J. Baynes, A. Weaver, A. Neale, T. Wade, M. Mehaffey et al. Development of the United States Environmental Protection Agency's Facilities Status Dashboard for the COVID-19 Pandemic: Approach and Challenges. *International Journal of Public Health*, 67:1604761, 2022. doi: 10.3389/ijph.2022.1604761 1

[8] R. Borgo, A. Abdul-Rahman, F. Mohamed, P. W. Grant, I. Reppa, L. Floridi et al. An Empirical Study on Using Visual Embellishments in Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2759–2768, 2012. doi: 10.1109/TVCG.2012.197 2

[9] M. A. Borkin, Z. Bylinskii, N. W. Kim, C. M. Bainbridge, C. S. Yeh, D. Borkin et al. Beyond Memorability: Visualization Recognition and Recall. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):519–528, 2016. doi: 10.1109/TVCG.2015.2467732 2

[10] M. A. Borkin, A. A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva et al. What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2306–2315, 2013. doi: 10.1109/tvcg.2013.234 2, 4

[11] Z. Bylinskii, M. A. Borkin, N. W. Kim, H. Pfister, and A. Oliva. Eye fixation metrics for large scale evaluation and comparison of information visualizations. In *Workshop on Eye Tracking and Visualization*, pp. 235–255. Springer, Cham, 2015. doi: 10.1007/978-3-319-47024-5_14 2

[12] C. Chen and Z. Liu. The state of the art in creating visualization corpora for automated chart analysis. In *Computer Graphics Forum*, vol. 42, pp. 449–470. Wiley Online Library, 2023. doi: 10.1111/cgf.14855 2

[13] J. Chen, M. Ling, R. Li, P. Isenberg, T. Isenberg, M. Sedlmair et al. Vis30k: A collection of figures and tables from ieee visualization conference publications. *IEEE Transactions on Visualization and Computer Graphics*, 27(9):3826–3833, 2021. doi: 10.1109/tvcg.2021.3054916 2

[14] X. Chen, W. Zeng, Y. Lin, H. M. Ai-Maneea, J. Roberts, and R. Chang. Composition and configuration patterns in multiple-view visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1514–1524, 2021. doi: 10.1109/TVCG.2020.3030338 2

[15] J. Choi, C. Oh, B. Suh, and N. W. Kim. Toward a unified framework for visualization design guidelines. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–7. ACM, NY, 2021. doi: 10.1145/3411763.3451702 9

[16] N. M. Cocoros, C. Kirby, B. Zambarano, A. Ochoa, K. Eberhardt, C. Rocchio, SB et al. RiskScape: a data visualization and aggregation platform for public health surveillance using routine electronic health record data. *American Journal of Public Health*, 111(2):269–276, 2021. doi: 10.2105/ajph.2020.305963 2, 3

[17] D. Deng, Y. Wu, X. Shu, J. Wu, S. Fu, W. Cui et al. VisImages: A Fine-Grained Expert-Annotated Visualization Dataset. *IEEE Transactions on Visualization and Computer Graphics*, 29(7):3298–3311, 2022. doi: 10.1109/TVCG.2022.3155440 2

[18] F. Elavsky, C. Bennett, and D. Moritz. How accessible is my visualization? Evaluating visualization accessibility with Chartability. In *Computer Graphics Forum*, vol. 41, pp. 57–70. Wiley Online Library, 2022. doi: 10.1111/cgf.14522 2

[19] A. J. Elliot. Color and psychological functioning: a review of theoretical and empirical work. *Frontiers in Psychology*, 6:127893, 2015. doi: 10.3389/fpsyg.2015.00368 7

[20] K. L. T. Fung, S. T. Perrault, and M. T. Gastner. Effectiveness of Area-to-Value Legends and Grid Lines in Contiguous Area Cartograms. *IEEE Transactions on Visualization and Computer Graphics*, 30(8):4631–4647, 2024. doi: 10.1109/TVCG.2023.3275925 2

[21] K. Hu, S. N. Gaikwad, M. Hulsebos, M. A. Bakker, E. Zgraggen, C. Hidalgo et al. VizNet: Towards a Large-Scale Visualization Learning and Benchmarking Repository. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–12. ACM, NY, 2019. doi: 10.1145/3290605.3300892 2

[22] S. C. S. Joyner, A. Riegelhuth, K. Garrity, Y.-S. Kim, and N. W. Kim. Visualization accessibility in the wild: Challenges faced by visualization designers. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–19. ACM, NY, 2022. doi: 10.1145/3491102.3517630 2, 9

[23] D. Jung, W. Kim, H. Song, J.-i. Hwang, B. Lee, B. Kim et al. Chartsense: Interactive data extraction from chart images. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 6706–6717. ACM, NY, 2017. doi: 10.1145/3025453.3025957 2

[24] N. W. Kim, G. Ataguba, S. C. Joyner, C. Zhao, and H. Im. Beyond Alternative Text and Tables: Comparative Analysis of Visualization Tools and Accessibility Methods. In *Computer Graphics Forum*, vol. 42, pp. 323–335. Wiley Online Library, 2023. doi: 10.1111/cgf.14833 2

[25] N. W. Kim, S. C. Joyner, A. Riegelhuth, and Y. Kim. Accessible visualization: Design space, opportunities, and challenges. In *Computer Graphics Forum*, vol. 40, pp. 173–188. Wiley Online Library, 2021. doi: 10.1111/cgf.14298 2

[26] W. B. Lammons, B. Moss, C. Bignell, C. Gale, A. MacBride, R. Ribas et al. Involving multiple stakeholders in assessing and reviewing a novel data visualisation tool for a national neonatal data asset. *BMJ Health & Care Informatics*, 30(1):e100694, 2023. doi: 10.1136/bmjhci-2022-100694 1

[27] A. Ledesma, M. Al-Musawi, and H. Nieminen. Health figures: an open source JavaScript library for health data visualization. *BMC medical informatics and decision making*, 16(1):38, 2016. doi: 10.1186/s12911-016-0275-6 2, 3

[28] C. Lee, T. Yang, G. D. Inchoco, G. M. Jones, and A. Satyanarayan. Viral Visualizations: How Coronavirus Skeptics Use Orthodox Data Practices to Promote Unorthodox Science Online. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–18. ACM, NY, 2021. doi: 10.1145/3411764.3445211 2

[29] P.-s. Lee, J. D. West, and B. Howe. Viziometrics: Analyzing visual information in the scientific literature. *IEEE Transactions on Big Data*, 4(1):117–129, 2017. doi: 10.1109/TBDATA.2017.2689038 2

[30] S. Lee, E. Kim, and K. A. Monsen. Public health nurse perceptions of Omaha System data visualization. *International Journal of Medical Informatics*, 84(10):826–834, 2015. doi: 10.1016/j.ijmedinf.2015.06.010 1

[31] L. Legenza, K. McNair, S. Gao, J. P. Lacy, B. J. Olson, T. R. Fritsche et al. A geospatial approach to identify patterns of antibiotic susceptibility at a neighborhood level in Wisconsin, United States. *Scientific Reports*, 13(1):7122, 2023. doi: 10.1038/s41598-023-33895-5 1

[32] M. Lisnic, C. Polychronis, A. Lex, and M. Kogan. Misleading Beyond Visual Tricks: How People Actually Lie with Charts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–21. ACM, NY, 2023. doi: 10.1145/3544548.3580910 2

[33] L. Y.-H. Lo, A. Gupta, K. Shigyo, A. Wu, E. Bertini, and H. Qu. Misinformed by Visualization: What Do We Learn From Misinformative Visualizations? 41(3):515–525, 2022. doi: 10.1111/cgf.14559 4

[34] O. G. McKinley, S. Pandey, and A. Ottley. Trustworthy by Design: The Viewer's Perspective on Trust in Data Visualization. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pp. 1–17. ACM, NY, 2025. doi: 10.1145/3706598.3713824 1, 2

[35] R. Mehta and R. Zhu. Blue or red? Exploring the effect of color on cognitive task performances. *Science*, 323(5918):1226–1229, 2009. doi: 10.1126/science.1169144 7

[36] K. Narayan and M. S. D. P. Nayak. Need for interactive data visualization in public health practice: examples from India. *International Journal of Preventive Medicine*, 12(1):16, 2021. doi: 10.4103/ijpvm.ijpvm_171_20 3

[37] M. A. Ofori, S. Lartey, P. Durneva, N. Jha, N. Mittal, S. Roy et al. Visual communication of public health data: a scoping review. *Frontiers in Digital Health*, 7:1555231, 2025. doi: 10.3389/fdgth.2025.1555231 2, 3

[38] A. Ottley. Consensus and Contradictions: A Cross-Organizational Analysis of Visualization Style Guides. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*, pp. 1–21. ACM, New York, 2026. doi: 10.1145/3772318.3791628 2, 9

[39] S. Pandey, O. G. McKinley, R. J. Crouser, and A. Ottley. Do you trust what you see? toward a multidimensional measure of trust in visualization. In *2023 IEEE Visualization and Visual Analytics (VIS)*, pp. 26–30. IEEE, 2023. doi: 10.1109/vis54172.2023.00014 1, 2

[40] S. Park, B. Bekemeier, A. Flaxman, and M. Schultz. Impact of data visualization on decision-making and its implications for public health practice: a systematic literature review. *Informatics for Health and Social Care*, 47(2):175–193, 2022. doi: 10.1080/17538157.2021.1982949 2, 3

[41] S. Park, B. Bekemeier, and A. D. Flaxman. Understanding data use and preference of data visualization for public health professionals: A qualitative study. *Public Health Nursing*, 38(4):531–541, 2021. doi: 10.1111/phn.12863 1, 2, 3

[42] J. Poco and J. Heer. Reverse-engineering visualizations: Recovering visual encodings from chart images. In *Computer Graphics Forum*, vol. 36, pp. 353–363. Wiley Online Library, 2017. doi: 10.1111/cgf.13193 2

[43] B. Preim and K. Lawonn. A Survey of Visual Analytics for Public Health. In *Computer Graphics Forum*, vol. 39, pp. 543–580, 2020. doi: 10.1111/cgf.13891 3

[44] J. Riddell, S. Letina, K. Skivington, D. Archambault, V. Wells, E. Long et al. Methods for interventions using networks to improve health: a narrative synthesis of methodological research on network data collection, visualisation and intervention. *Social Networks*, 84:202–219, 2025. doi: 10.1016/j.socnet.2025.10.003 3

[45] Y. Senathirajah, D. R. Kaufman, K. D. Cato, E. M. Borycki, J. A. Fawcett, and A. W. Kushniruk. Characterizing and Visualizing Display and Task Fragmentation in the Electronic Health Record: Mixed Methods Design. *JMIR Human Factors*, 7(4):e18484, 2020. doi: 10.2196/18484 1

[46] A. Shaban-Nejad, M. Lavigne, A. Okhmatovskaia, and D. L. Buckeridge. PopHR: a knowledge-based platform to support integration, analysis, and visualization of population health data. *Annals of the New York Academy of Sciences*, 1387(1):44–53, 2017. doi: 10.1111/nyas.13271 2

[47] V. Sharma, O. Cassetti, L. Winning, M. O'Sullivan, and M. Crowe. Protocol for developing a dashboard for interactive cohort analysis of oral health-related data. *BMC Oral Health*, 23(1):238, 2023. doi: 10.1186/s12903-023-02895-2 3

[48] D. Skau, L. Harrison, and R. Kosara. An evaluation of the impact of visual embellishments in bar charts. In *Computer Graphics Forum*, vol. 34, pp. 221–230. Wiley Online Library, 2015. doi: 10.1111/cgf.12634 2

[49] H. Song, A. Cho, C. X. Bearfield, and J. Stasko. Visualizing trust: How chart embellishments influence perceptions of credibility. *IEEE Transactions on Visualization and Computer Graphics*, 32(1):1306–1316, 2026. doi: 10.1109/TVCG.2025.3634785 2

[50] L. Su, A. P. Cui, and M. F. Walsh. Trustworthy blue or untrustworthy red: The influence of colors on trust. *Journal of Marketing Theory and Practice*, 27(3):269–281, 2019. doi: 10.1080/10696679.2019.1616560 7

[51] P. S. Sullivan, C. Woodyatt, C. Koski, E. Pembleton, P. McGuinness, J. Taussig et al. A data visualization and dissemination resource to support HIV prevention and care at the local level: analysis and uses of the AIDSVu public data resource. *Journal of Medical Internet Research*, 22(10):e23173, 2020. doi: 10.2196/23173 1, 2, 3

[52] X. Tan, X. Suo, W. Li, L. Bi, and F. Yao. Data visualization in healthcare and medicine: a survey. *The Visual Computer*, 41(5):3037–3058, 2025. doi: 10.1007/s00371-024-03586-x 3

[53] E. R. Tufte and P. R. Graves-Morris. *The visual display of quantitative information*, vol. 2. Graphics press Cheshire, CT, 1983. 2

[54] R. O. Valdiserri and P. S. Sullivan. Data visualization promotes sound public health practice: the AIDSvu example. *AIDS Education and Prevention*, 30(1):26–34, 2018. doi: 10.1521/aeap.2018.30.1.26 2, 3

[55] V. L. West, D. Borland, and W. E. Hammond. Innovative information visualization of electronic health record data: a systematic review. *Journal of the American Medical Informatics Association*, 22(2):330–339, 2015. doi: 10.1136/amiajnl-2014-002955 3

[56] B. L. Wimer, L. South, K. Wu, D. A. Szafir, M. A. Borkin, and R. A. Metoyer. Beyond Vision Impairments: Redefining the Scope of Accessible Data Representations. *IEEE Transactions on Visualization and Computer Graphics*, 30(12):7619–7636, 2024. doi: 10.1109/TVCG.2024.3356566 2

[57] M. Zakkar and K. Sedig. Interactive visualization of public health indicators to support policymaking: An exploratory study. *Online Journal of Public Health Informatics*, 9(2):e190, 2017. doi: 10.5210/ojphi.v9i2.8000 1

[58] Y. Zhang, Y. Sun, L. Padilla, S. Barua, E. Bertini, and A. G. Parker. Mapping the Landscape of COVID-19 Crisis Visualizations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–23. ACM, NY, 2021. doi: 10.1145/3411764.3445381 2

A Sources That We Selected Public Health Organizations

Table 2: The complete list of sources (with their URLs) we used to select public health organizations for our study.

Source	URL
2025 Forbes Top Charities	https://www.forbes.com/lists/top-charities/
NIH Institutes and Centers	https://www.nih.gov/institutes-nih/list-institutes-centers
HHS Agencies and Offices	https://www.hhs.gov/about/agencies/hhs-agencies-and-offices/index.html
UN and Multilateral Organizations	https://www.un.org/en/about-us/un-system
University of Illinois Chicago, School of Public Health	https://publichealth.uic.edu/global-health-program/applied-practice-experience/un-multilateral-organizations/

B Finalized Codebook

Table 3: Description of the finalized codebook with 6 major categories and the variables nested in each. Data types of the variables are in parentheses. *Assignment* refers to how the variable was designated.

Theme	Short Description	Variables (with Data Type)	Assignment
Chart Family	High-level chart type	Chart Type (Single Option): Bar Chart, Area Chart, Line Chart, Scatter Plot, Map, Table, Pie Chart, Histogram, Box Plot, Heatmap, Flow Diagram, Network Variable, Tree Diagram, Timeline, Dashboard, Infographic, Other	GPT
Data Context	Data task	Data Context (Single Option): Temporal Trend, Geographic, Categorical Comparison, Distribution, Composition Part to Whole, Ranking Ordered, Correlation Relationship, Flow Process, Hierarchical, Network Relational, Multivariate, Text Narrative, Mixed	GPT
Structural Features	Basic structure elements	Direct Labels (Boolean), Legend (Boolean), Grid Lines (Boolean), Title (Boolean), Y-Axis (Boolean), X-Axis (Boolean), Labeled Axes (Boolean), Multiple Panels (Boolean), Missing Data (Boolean)	GPT
Visual Embellishments	Decorative elements	Iconography (Boolean), Photograph (Boolean), Medical Imaging (Boolean), Unique Shapes or Lines (Boolean), Text Annotations (Boolean), X-Axis (Boolean), Text Annotations (Boolean)	GPT
Color	Color palette	Color Blind Risky (Boolean), Grayscale (Boolean), Palette Size (Number), HTML HEX Codes (List), HSV Color Names (List)	GPT, Measurement
Text	Text inside the chart	Visualization Text (Text), Text Density (Number)	OCR, Measurement

C Preliminary Meta Data Analysis

Along with the scraping of visualizations, we collected web metadata fields including Last-Modified Date, Title, Author, and Description. We considered analyzing the timeline of visualizations produced and performing topic modeling. In this section, we discuss our decisions and preliminary findings. We encourage future work to explore the metadata further in their analyses.

C.1 Timeline

The Last-Modified date was collected from the HTTP header response of the webpage for the scraped image. The Last-Modified date may not be a reliable metric given the development of the visualizations in our corpus. Approximately 50% of the visualizations (n = 2,207) have a Last-Modified date. A webpage with dynamic content, e.g., scripts executed on request, can update the Last-Modified date to the date the visualization was collected (n = 426). The Last-Modified date may not reflect the actual timeline when a visualization was created; we have observed that organizations reuse many visualizations, as further demonstrated by the removal of duplicate images in Section 3.4. We did not proceed with a timeline analysis of the collected visualizations; however, for completeness, we retained the Last-Modified date field in the public dataset.

C.2 Topic Modeling

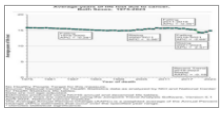
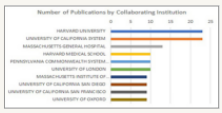
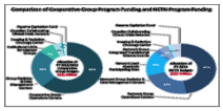
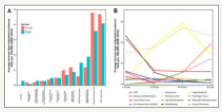
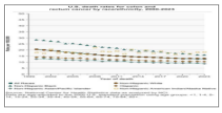
We collected Title, Author, and Description from the meta tags of the HTML content related to the webpage of a visualization. Description and Author are applied unevenly across the visualization corpus: approximately 73% of visualizations have descriptions (n = 3,122), and less than 2% have a specified author (n = 65). All visualization webpages have titles; therefore, we continued with topic modeling of the title field to ensure an even representation of the corpus. We had difficulty finding meaningful insights from the topic groups, which we suspect is due to two potential reasons. 1) Titles can be pretty short, so topic models can easily group items on trivial words. It would take fine-tuning and editing of the titles to potentially garner meaningful insights. 2) Many of the public health organizations in our study have a specific purpose or cause, such as cancer, smoking, and mental health. Therefore, in our preliminary analysis, we identified topic groups based on the organizations' specific cause. For

instance, smoking, drugs, and drinking were all one topic group. Ultimately, we found these groupings inconsequential given our primary focus on overall design patterns in the public health space.

D EXAMPLES OF PALETTE ACCESSIBILITY UNDER SIMULATED COLOR VISION DEFICIENCY

Example palette behavior under simulated color vision deficiency

Each cell shows the minimum pairwise CIEDE2000 color difference (ΔE00). When a risky pair is present, the lowest-distance pair is shown as swatches.

Visualization	Normal	Protanomaly	Deuteranomaly	Tritanomaly
	min ΔE00 = 7.17 **No risky pair**	min ΔE00 = 7.17 **No risky pair**	min ΔE00 = 7.17 **No risky pair**	min ΔE00 = 7.17 **No risky pair**
	min ΔE00 = 21.57 **No risky pair**	min ΔE00 = 1.25 **1.25** #ff7f0e vs #2ca02c	min ΔE00 = 3.33 **3.33** #ff7f0e vs #bcbd22	min ΔE00 = 12.03 **No risky pair**
	min ΔE00 = 9.20 **No risky pair**	min ΔE00 = 5.70 **No risky pair**	min ΔE00 = 6.95 **No risky pair**	min ΔE00 = 4.27 **4.27** #0f5760 vs #214a71
	min ΔE00 = 11.13 **No risky pair**	min ΔE00 = 3.50 **3.50** #F2716E vs #1B9E77	min ΔE00 = 4.17 **4.17** #984EA3 vs #377EB8	min ΔE00 = 2.75 **2.75** #F2716E vs #F46D43
	min ΔE00 = 5.51 **No risky pair**	min ΔE00 = 5.77 **No risky pair**	min ΔE00 = 5.14 **No risky pair**	min ΔE00 = 5.71 **No risky pair**

Examples illustrate both stable palettes and condition-specific collapses under simulated CVD.

Figure 5: Representative examples from the corpus illustrating palette behavior under normal vision and three simulated color vision deficiency (CVD) conditions: protanomaly, deuteranomaly, and tritanomaly. Each cell reports the minimum pairwise perceptual color difference using CIEDE2000 (ΔE_{00}). When a risky color pair is detected, the lowest-distance pair is shown as color swatches and annotated with its corresponding ΔE_{00} value. These examples illustrate both stable palettes and condition-specific color collapses that may reduce distinguishability under simulated CVD.

E Visualization Examples from Design Profiles

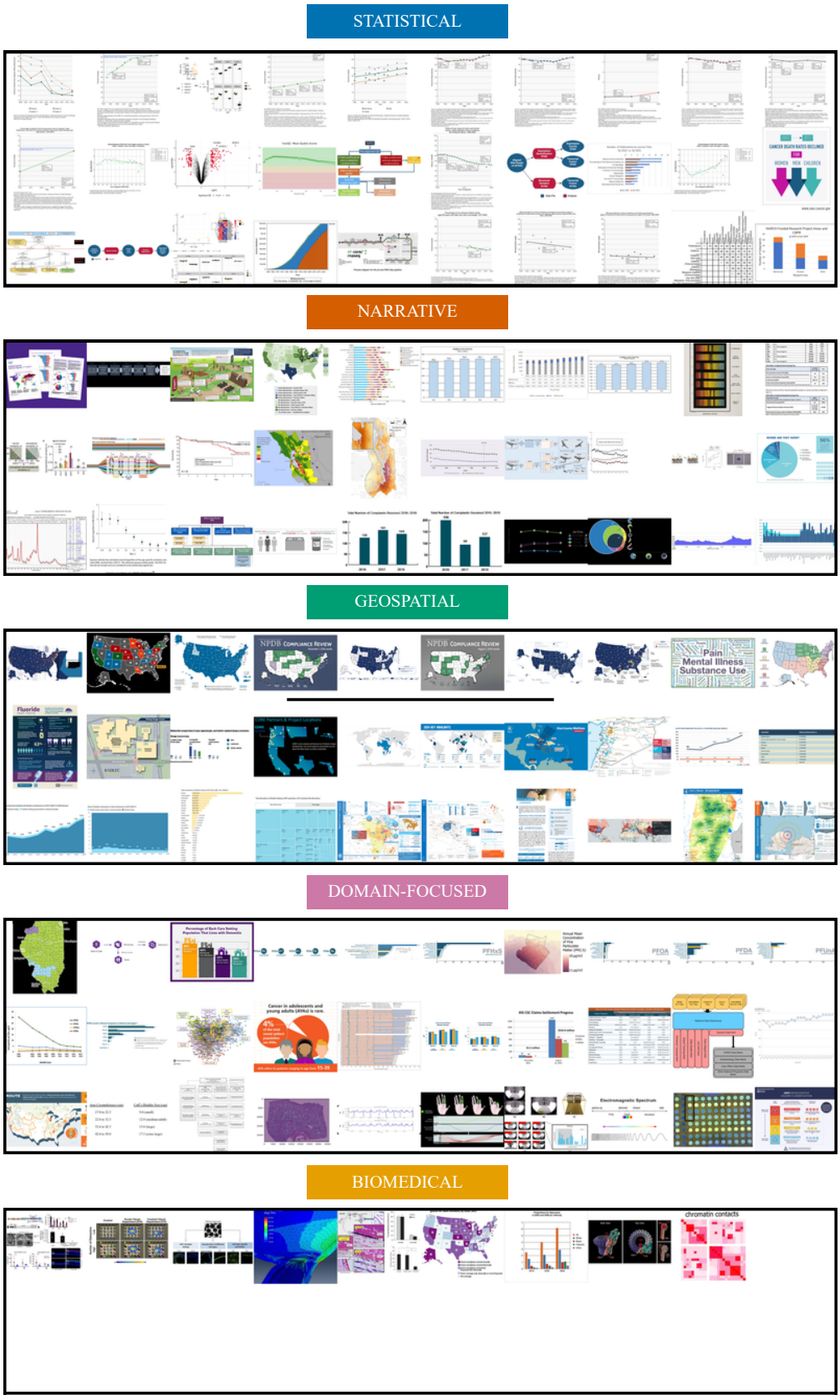


Figure 6: Sampled visualizations of all five design profiles.