Consensus and Contradictions: A Cross-Organizational Analysis of Visualization Style Guides

Alvitta Ottley
Department of Computer Science and Engineering
Washington University in St. Louis
St. Louis, Missouri, USA
alvitta@wustl.edu

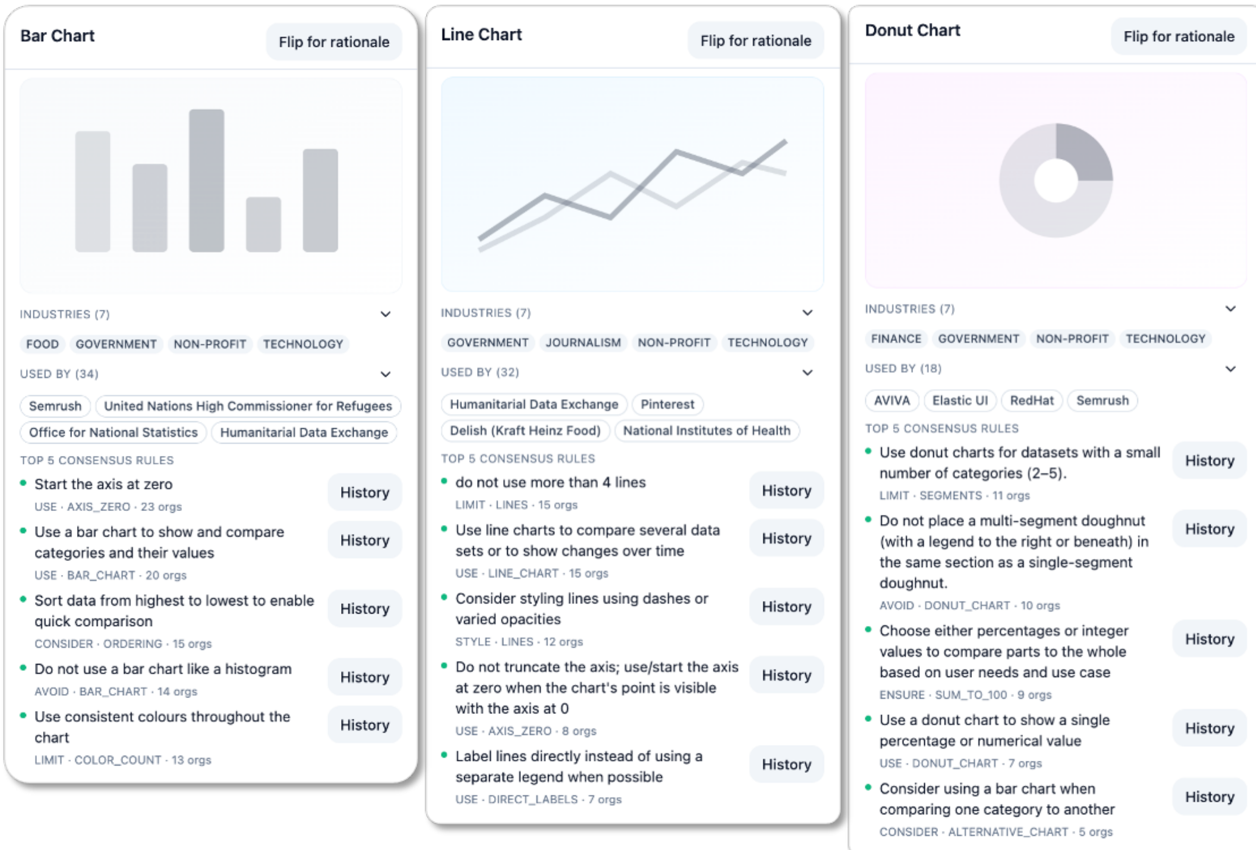


Figure 1: Flash cards summarizing design guidelines that were recommended by at least three organizations from at least three distinct industries for bar charts, line charts, and donut charts.

Abstract

Should you place three pie charts side by side, or should you avoid pie charts altogether? Publicly available visualization style guides offer contradictory answers to such questions. Despite their growing influence on how people encounter data, these guides are seldom studied as a collective phenomenon. Addressing this gap, this paper presents the first systematic analysis of 53 publicly accessible visualization style guides from diverse domains, including journalism, government, non-profit, corporate, and academic sectors. We build a standardized corpus, conduct a multi-method analysis that reveals both consensus and contradiction, and develop a companion *Guidelines Explorer* to support transparency and future use. This work sheds light on organizational visualization design norms and provides a foundation for future work that helps bridge the gap between academic and industry practices. In doing so, we help reframe style guides as sociotechnical artifacts that encode values as much as design rules.

CHI '26, Barcelona, Spain

ACM ISBN 979-8-4007-2278-3/26/04
https://doi.org/10.1145/3772318.3791628

CCS Concepts

• **Human-centered computing → Visualization design and evaluation methods**; **Visualization theory, concepts and paradigms**.

Keywords

Data Visualization, Guidelines, Style Guides, Design Values

ACM Reference Format:
Alvitta Ottley. 2026. Consensus and Contradictions: A Cross-Organizational Analysis of Visualization Style Guides. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 21 pages. https://doi.org/10.1145/3772318.3791628

1 Introduction

One style guide states, "if needed, up to three pie charts can be placed side by side" [AVV-FINANCE]. In another, they contradict by saying "do not use multiple pie chart[s] for comparison. Single pie charts are already difficult to read, ..."[HDX-NONPROFIT]. Both cannot be true, and yet both are circulated as authoritative advice. Such contradictions are not isolated cases; practitioners have long noted these inconsistencies anecdotally, but we lack systematic evidence about how often they occur, where they appear, and what patterns they follow. As governments, news outlets, health agencies, and visualization enthusiasts increasingly rely on style guides to design dashboards and charts for the public, these documents play a growing role in shaping how society encounters data [49]. Yet, despite this influence, visualization style guides have rarely been studied as a collective phenomenon [12, 18]. What exactly do they recommend? How consistent are they across organizations and industries? And what *values* and assumptions do they embed?

These questions matter because guidelines are inherently contextual. What works for a newsroom may not suit a public health agency, and different communities (e.g., journalists, designers, or researchers) prioritize different values when judging what counts as a "good" visualization. For example, academic work often centers on perceptual accuracy and efficiency [14, 19, 26], while practitioners may have to balance these concerns with broader goals such as sustaining engagement, reinforcing trust, or emphasizing a narrative takeaway. Even when efficiency is the focus for both groups, strategies may differ [44, 45]. In controlled studies, researchers may minimize visual embellishment to isolate the data signal, while practitioners might improve efficiency by simplifying datasets or highlighting key messages to enhance accessibility. Importantly, perspectives are shifting. The field has moved beyond historically minimalist design, and researchers today are increasingly open to studying how aesthetics, annotation, and even deliberate embellishment can enrich communication [2, 7, 11, 53, 60]. This growing body of work shows that visual embellishments can improve memorability, engagement, or affect, suggesting that accounting for these effects leads to more nuanced perceptual studies and more balanced guidance for practice.

Additionally, researchers have increasingly turned their attention to studying visualization practice itself, examining how professionals design, justify, and evaluate their work. Scholars have interviewed practitioners to document real-world workflows and best practices [3, 6, 43, 45, 48, 50], while other studies interrogate dimensions that practitioners care about directly, including what makes visualizations "misleading,", "aesthetic," "trustworthy," or "memorable" [4, 25, 31, 34, 36]. Grassroots initiatives, such as the Data Visualization Style Guide project [49], are cataloging organizational practices. Their effort has highlighted the difficulties of studying guidelines on a large scale. This is because the formats of these documents vary considerably, and they tend to adopt inconsistent naming conventions and terminology. These challenges have led to calls for establishing standardized formats and systematic methods for collecting, analyzing, and comparing visualization guidelines [12, 18].

To address this gap, this paper presents a systematic analysis of 53 publicly accessible style guides from journalism, government, non-profit, corporate, and academic organizations. We extracted and transformed their heterogeneous recommendations into a standardized corpus of 2,120 chart-specific guidelines. Using a hybrid human–AI pipeline, we extracted action–target units, clustered them into 226 unique guideline patterns. We analyzed these guidelines both quantitatively and qualitatively. Our analysis reveals both consensus and conflict. Across all industries, bar chart guidance dominates, suggesting a strong consensus among organizations that bar charts are the "safe" default for public communication. We find broad agreement on practices such as direct labeling and limiting data points in charts, but disagreements remain over using pie charts and the necessity of a zero baseline on the y-axis.

To support exploration of this corpus, we developed the *Guidelines Explorer* website, which presents the guidelines in an interactive, searchable interface. This dataset and platform enable, for the first time, researchers, practitioners, and educators to critically engage with industry guidelines and systematically identify consensus practices, contested rules, and industry-level emphases. In doing so, we aim to mitigate uncritical adoption, support the teaching of "best practices," and encourage further research into not just what organizations recommend, but how those recommendations encode institutional values and priorities. These are reflected in our contributions:

- First, we curate and release a standardized, cross-organizational corpus of 53 visualization style guides, representing journalism, government, health, corporate, and academic domains.
- Second, we provide a multi-method analysis that surfaces both consensus (e.g., common practices around bar and line charts) and contradictions (e.g., contested use of pie charts and axis baselines), reframing guidelines as value-laden sociotechnical artifacts rather than universal rules.
- Third, we introduce the Guidelines Explorer as a companion tool that makes the corpus publicly accessible and explorable at https://washuvis.github.io/styleguides/.

2 Background

A rich body of research underpins today's visualization conventions, grounded in decades of work across perception, cognition, information design, and HCI. Foundational contributions such as Bertin's system of graphical primitives [5] and Card et al.'s data-mapping framework [9] offered some of the first systematic ways of thinking about how visual encodings convey meaning. Building on this, Tufte's principles for reducing visual clutter [55], empirical studies by Cleveland and McGill [14], and perceptual syntheses such as Ware's [58] clarified which encodings people judge most accurately and why. As interactive visualization matured, Shneiderman's information-seeking mantra [51] reframed design around user goals and exploratory workflows.

More recent work has examined how factors such as cognitive state, chart complexity, narrative structure, annotation, interactivity, and text placement influence comprehension and decision-making (e.g., [13, 29, 32, 46, 53, 61]). This research on perception and cognition offers a solid foundation for understanding how visualizations convey information, and these principles are now embedded, explicitly or implicitly, into modern visualization tools and recommendation engines such as Vega-Lite [47], CompassQL [59], Draco [38, 62, 63], and Tableau [54]. Moreover, organizations that regularly produce visualizations often distill these accumulated research findings into practical guidance, formalized as *style guides*. These documents summarize decades of visualization research and practitioner insights into guidelines and examples that enhance clarity, consistency, and communication across various fields and audiences [12, 18].

2.1 Prior Research on Visualization Guidelines

Style guides have long been used to codify design practices in HCI and Data Visualization, offering rules and recommendations that shape how artifacts are created, shared, and interpreted [1, 8, 23, 39]. One of the earliest and most influential examples is Brinton's *Graphic Methods for Presenting Facts* (1919), often credited as the first visualization style guide [8]. Brinton's work combined general guidance with checklists and explicit "rules for graphic presentation," ranging from layout (e.g., "The general arrangement of a chart should proceed from left to right") to textual framing (e.g., "Make the title of a chart so complete and so clear that misinterpretation will be impossible"). This rule-based rhetoric persists today. Modern organizational style guides continue to present concrete dos and don'ts, often supported by examples, templates, and rationales, to promote clarity and maintain consistency across teams [12, 18].

A handful of academic studies have examined style guides or guidelines directly. Kandogan and Lee [30] used a grounded-theory approach to analyze 550 guidelines drawn from 24 sources, including academic publications, books, and blogs, highlighting that guidance often spans data, tasks, user expertise, and insight; not just visual form. Elder and Cesal [18] took a more practice-centered view and analyzed eight style guides from nonprofit, government, and corporate organizations, comparing their contents to Brinton's original rules. They observed that many contemporary guides still echo Brinton's recommendations, even as they expand coverage of chart types and conventions.

Diehl et al. [17] examined guidance on VisGuides, a public forum for visualization advice, and found that guidelines were mentioned in about 41.5% of practitioner conversations, with many related to color. They noted that guidelines were less prevalent than expected, which may indicate the difficulty of applying them to practical problems. Choi et al. [12] studied 226 individual guidelines from 19 online sources and proposed that guidelines can typically be decomposed into eight components (purpose, problem, approach, alternatives, chart type, chart elements, task type, and data type). Categorizing guidelines along these dimensions, they found that readability, credibility, accuracy, clarity, and accessibility emerged as common purposes. Their structure suggests opportunities for automating guideline creation and analysis.

Taken together, these studies illuminate important facets of visualization guidance, such as academic recommendations, conceptual structure, practitioner discourse, and small-scale organizational practice. However, each examines only a narrow portion of the broader ecosystem. They do not provide a unified, cross-organizational perspective on how style guides differ across sectors, where recommendations converge or conflict, or how institutional values shape their content. These limitations motivate the need for a more comprehensive analysis.

2.2 Challenges in Studying Guidelines

A central challenge is inconsistency, as guidelines frequently contradict one another. In practice, these contradictions are widely recognized anecdotally but have not been systematically documented at scale. Practitioners often rely on situated expertise or organizational norms, rather than formal research-based recommendations, while researchers produce empirically grounded advice that may be difficult to translate into practice [43]. This disconnect obscures which guidelines are broadly adopted, which are contested, and which reflect the particular priorities of specific organizations. Additionally, with the advocacy for *Value Sensitive Design* [10, 20, 21], ethics [15, 16, 28, 29, 57], and other social constructs in visualization practice suggest that design conventions may intentionally embody sociotechnical values, stemming from institutional goals, professional norms, tooling constraints, and assumptions about audiences. Understanding style guides, therefore, requires not only cataloging their rules but also examining the contexts in which those rules take shape.

Altogether, the landscape of visualization guidelines remains fragmented. What is missing is a systematic, cross-organizational analysis that reveals how style guides align, diverge, and encode institutional values. Our work addresses this gap by analyzing a corpus of 53 style guides spanning journalism, government, nonprofit, academic, and corporate sectors, and by introducing the *Guidelines Explorer* as a tool for transparent, comparative analysis.

3 Data Collection & Standardization

1 We began with a list of style guides from the Data Visualization Style Guide website[1], which hosts a grassroots dataset of style guides across various sectors. At the time of access, there were 78 entries. The collection was stored as a *Google Sheet*, documenting the timestamp when each record was added, the company name, and the URL of the style guide document. From this list, we selected only those style guides that were publicly accessible online. Several were excluded when links were broken or content was unavailable. There were also duplicate entries. This resulted in a final set of 53 style guides.

2 Each used a different format, ranging from neatly structured PDFs to loosely formatted web pages. As a result, data extraction

[1](https://www.datavizstyleguide.com/)

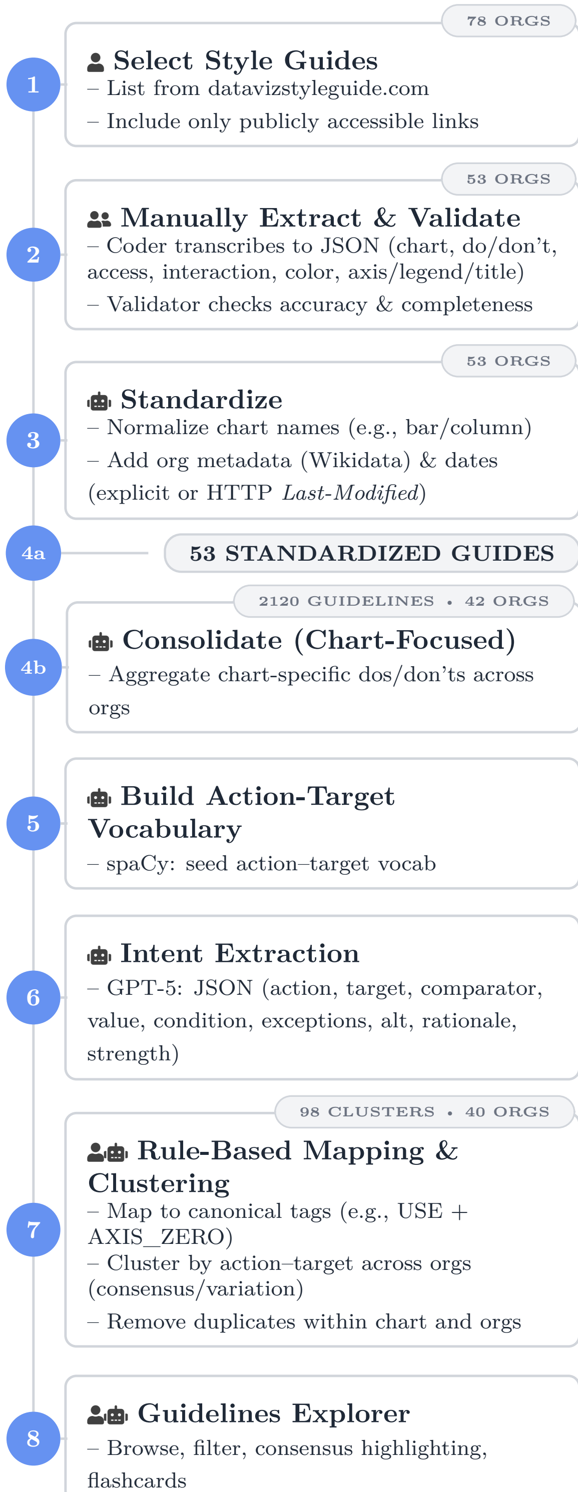


Figure 2: The workflow for collecting, standardizing, consolidating, and clustering 78 style guides. We use 🤖 and 👤 to distinguish between automation and human effort.

required careful human effort. We adopted a two-step process, with a coder and a validator assigned to each document. Four volunteer coders, each with visualization experience and interested in using the dataset for their own projects, reviewed the source material and then pasted the relevant content into structured JSON. The JSON schema organizes information around organizational metadata, chart types, accessibility notes, color choices, and recommendations for elements such as axes, tooltips, and legends (see Appendix A for schema details). Validators then checked each JSON file against the original document to ensure completeness, especially when content appeared across multiple webpages or in unstructured text.

3 Once collected, the dataset was cleaned and standardized. A central challenge was inconsistency in naming conventions: the same chart type often appeared under multiple terms (e.g., "bar chart", "bar graph",," "column chart," and "vertical bar chart"). We also enriched the data with organizational metadata. Each organization was assigned an ORG_ID (based on an abbreviation) and an industry label using the Wikidata API. When multiple industry categories were returned, we used the first listed; for example, *Apple* is associated with various categories, such as *electronics*, *mobile phone industry*, *digital distribution*, and *software development*. The industry assigned was *electronics*, later clustered as *technology* for the analysis in Sections 6 and 7. To capture temporal context, we logged the creation or last-updated date of each guide. In some cases, this was straightforward, such as when an explicit date was provided on a PDF or webpage, or when the guide was hosted in a GitHub repository. In other cases, we relied on HTTP headers to retrieve the Last-Modified tag. All dates and metadata were normalized into a standard vocabulary, and a consistent schema was applied across all guides.

4a The result is a collection of standardized JSON files representing 53 organizational style guides. 4b In addition to the individual JSON files, we created a consolidated, chart-focused dataset that aggregates all dos and don'ts across the corpus, enabling systematic comparison of guidelines at both the organizational and cross-organizational levels.

4 Style Guides Structure

We analyzed content in each major component of our standardized JSON schema to identify emerging patterns when considering the guides as complete documents (see Appendix A for schema details). We categorized components as absent (0 items), minimal (1–2 items), or substantial (3+ items). This thresholding provides an interpretable indicator of structural emphasis without making assumptions about writing style or document length.

4.1 Corpus-Level Structural Patterns

Across the corpus, we noted different levels of emphasis on the components of visualization practice. We can observe in Figure 3 that charts (77% substantial) and color guidelines (70%) are the only components that are consistently detailed. General task guidance appears rarely (4% substantial), but tasks are typically specified within chart-type recommendations rather than as standalone guidance. When measured at this chart-level granularity, 66% of guides include substantial task-type references.

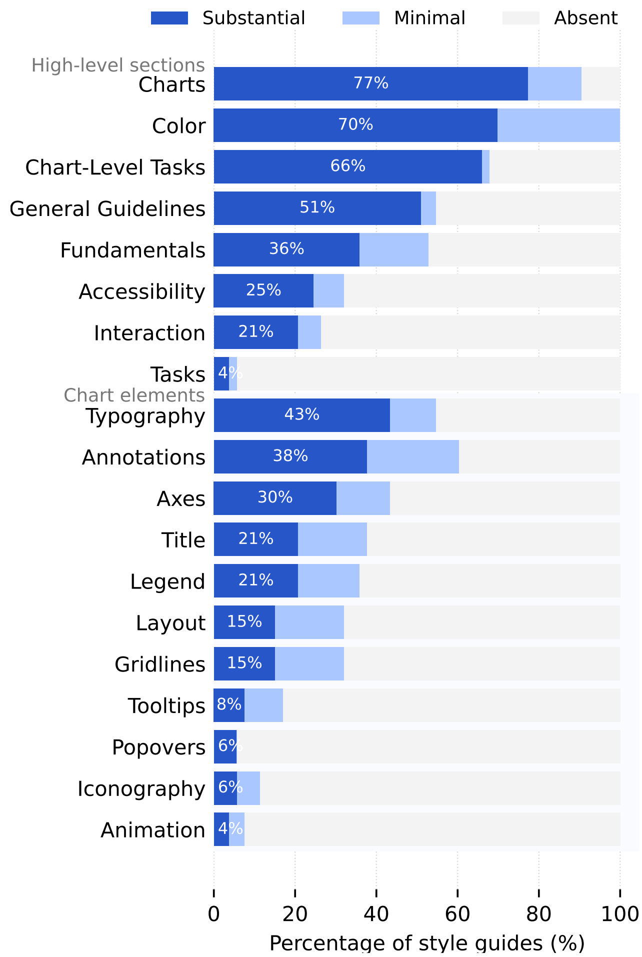


Figure 3: Distribution of substantial, minimal, and absent content across style-guide components. Most guides provide detailed recommendations for charts, color, and chart-level tasks, but offer far less guidance on dynamic or secondary chart elements.

Chart-element components exhibit similar variation. Typography (43% substantial) and annotations (38%) are more frequently developed, whereas interaction- and motion-related elements, such as animation, tooltips, and popovers, are substantial in only 4–8% of guides. These results show that most style guides prioritize static and structural aspects of visualization design over interactive or dynamic behaviors.

Structural completeness also varies widely. Twenty guides (37.7%) contain four or fewer substantially populated components, while five (9.4%) contain ten or more. The majority (52.8%) had 5–9 substantial components. Because our analysis is based on a schema that represents a superset of potential components, these counts should be interpreted as differences in emphasis rather than omissions. Altogether, these patterns show how organizations allocate attention to visualization design and motivate our next analysis, which examines variation across organizations and industries.

4.2 Structural Differences Within and Across Industries

Figure 4 highlights clear structural differences across industries. Government guides show the broadest substantial coverage, with most including charts (89%), chart-level tasks (89%), general guidelines (67%), and color (78%). Technology guides cover many of the same areas, including charts (77%), chart-level tasks[2] (77%), and color (73%), but are more uneven, offering relatively little coverage of accessibility (23%) and chart-element components such as typography (14%). Journalism shows a much narrower profile: while two-thirds include substantial chart and color guidance (67% each), few address accessibility (17%) or interaction (17%), and chart-level tasks appear only sporadically (17%). Non-profit guides emphasize charts (78%) and often omit accessibility guidance (11%). Across all industries, the matrix also reveals noticeable within-group variation, indicating that style guides do not follow a single structural template. Organizations differ considerably in which components they emphasize, resulting in diverse structural profiles.

While structural comparisons reveal which topics organizations prioritize, they do not reveal the nature of the instructions themselves. Two documents may devote equal space to charts yet offer very different recommendations. We therefore shift from the structure of style guides to the rules they prescribe, analyzing their intents and the underlying value framings they embody.

5 Intent Extraction

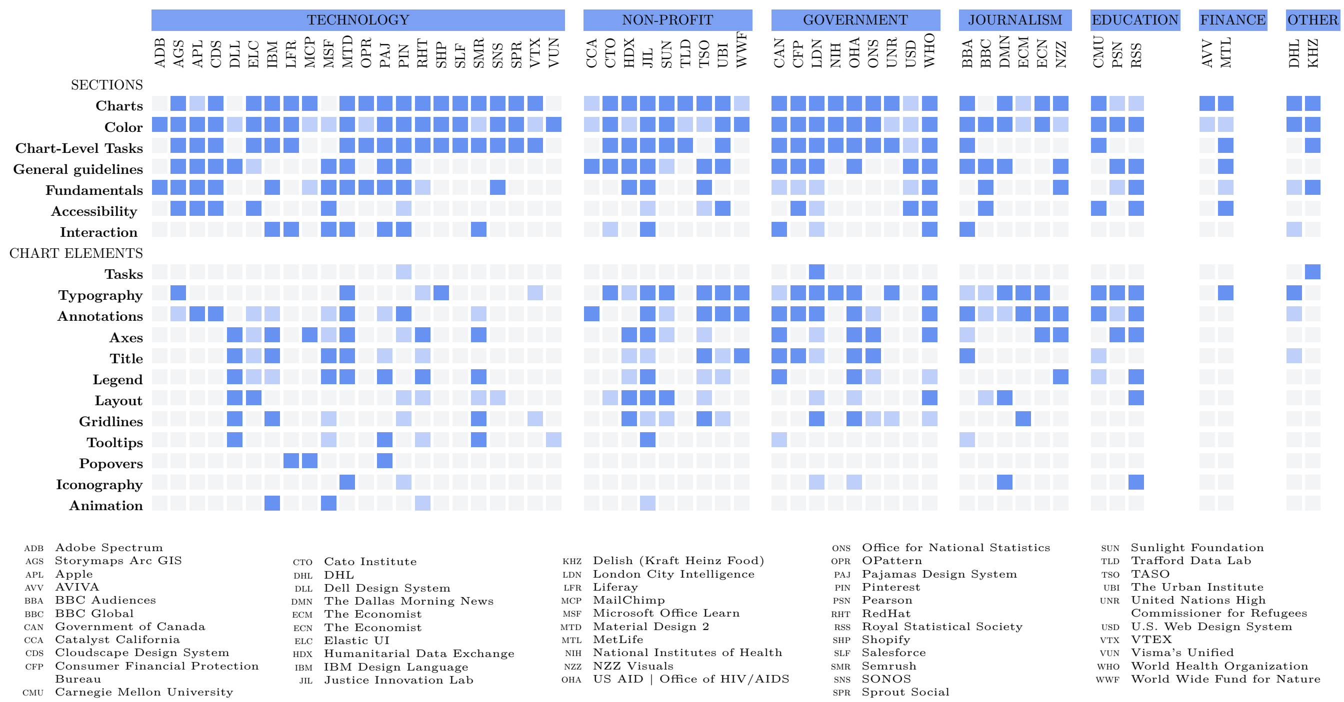
Although our corpus also includes guidance on colors, accessibility, interaction, and chart elements such as titles, legends, and tooltips, we restrict the following analysis to *chart-specific recommendations*. We made this choice to enable a focused and actionable discussion. Chart selection and design often serve as a gateway to other design considerations, providing the clearest entry point for addressing our research questions about consensus and contradictions.

5.1 Constructing a Controlled Action–Target Vocabulary from Heterogeneous Guidelines

The JSON format standardized our corpus, but the individual guidelines were free-text copied or transcribed from the original documents and were highly heterogeneous. Some were rigid prescriptions (e.g., "👎 [Don't] rotate bar labels" [vtx-technology]), while others were softer suggestions (e.g., "Consider using a horizontal histogram when range names are long, or there are many ranges" [smr-marketing]). No single extraction method would reliably capture this diversity. Rule-based approaches were brittle, and LLMs tended to over-generalize or infer unstated details. We therefore designed a three-step hybrid human–AI pipeline that combines automation with controlled refinement, aiming to capture for each guideline

[2]Tasks vs. Chart-level tasks. We distinguish (1) Tasks: a list of task concepts defined by the guide (stored as objects in tasks, often with only name and optional definition), from (2) Chart-level tasks: an explicit mapping from task to recommended chart types (captured when charts[i].tasks is present and non-empty). In other words, chart-level tasks are a property of charts; they exist only when the guide provides an explicit chart–task association.

Figure 4: Rows correspond to major sections of the style guides, with columns for organizations grouped by industry. We use ■ for substantial content, ■ for minimal content, and ■ when the section is absent.



the referenced element, the prescribed action, its rationale, and any specified exceptions.

5 To create our "alphabet," we enumerated the linguistic space of possible guideline "actions" and "targets." We used *spaCy* [27], an open-source NLP library, to automatically extract verb lemmas and noun phrases from all guideline text. This yielded a large set of raw candidates (534 unique action strings and 531 unique target strings), including variants (e.g., "category" and "more categories") and stop words (e.g., "it" and "let"). To reduce this set, we removed stop words and normalized each string by stripping punctuation and converting it to a canonical form (e.g., converting "axis labels" to **AXIS_LABELS** and "label directly" to **DIRECT_LABELS**). We then merged semantically equivalent expressions (e.g., "bar chart," "column chart," etc.), consistent with *Step 3*. After applying these consolidation and pruning steps, our initial alphabet was the 60 most frequent actions and targets.

To refine the automated enumeration and help identify conceptually important but infrequent items, a researcher manually tagged a random sample of 100 action–target pairs, using the most common mined terms as seeds. This hand-coding step helped identify overly fragmented items, confirmed which noun phrases should collapse into a single target (e.g., chart types were grouped under **CHART_TYPE** or **ALTERNATIVE_CHART** depending on the context), and surfaced synonyms among actions (e.g., "plot," "display," and "visualize"). These refinements narrowed the alphabet into a clean, controlled set of standardized actions and targets, which was then used as a constraint for the *GPT-5* extraction.

5.2 Applying the Action–Target Vocabulary via LLM Extraction

6 We used *GPT-5* [40], a large language model accessible through an API, to enrich these action–target pairs with additional structure[3] We designed a detailed prompt instructing the model to convert each guideline into a structured JSON format with fields for `action`, `target`, `comparator`, `units`, and `values` from a given set and extract `condition`, `exceptions`, `alternatives`, `rationale`, `strength`, and `plain intent`. All fields had to be explicitly stated in the guideline text, except `strength`, which the model inferred from linguistic cues such as "always," "must," or "may."

To evaluate the reliability of this extraction, we compared the 100 coded guidelines described in *Step 6*, with the LLM labels for action and target. In the first round, raw agreement was moderate, yielding a Cohen's $\kappa \approx .62$. Inspection showed many mismatches stemmed from action synonyms or target differences arising from phrasing. We therefore introduced a synonym map for normalizing actions (e.g., **USE**: [**USE**, **ADD**, **SET**, **PLACE**, **CONSIDER**, **APPLY**, **SHOW**, **ENSURE**]) and targets (e.g., **COLOR**: [**COLOR**, **COLOR_**

[3]We made all API call with temperature = 0 and top_p = 1.0 to minimize randomness, max_tokens = 800 to capture full JSON outputs. We used the default frequency and presence penalties. Prompts were provided exactly as specified in the supplementary material.

PALETTE, COLOR_COUNT]). After applying this normalization, the second-round Cohen's $\kappa \approx .84$ reflected strong alignment between human and LLM. We did not force perfect matches, as the remaining disagreements predominantly involved equally defensible interpretations, such as negative vs. positive framing for action–target pairs (e.g., USE•SINGLE_COLOR vs. LIMIT•COLOR) or compound recommendations where either tag was reasonable.

5.3 Annotating and Validating Values in Guideline Rationales

To analyze the rationales underlying visualization guidelines, we annotated a subset of guidelines with value tags, each corresponding to a normative design value (e.g., `clarity`, `accessibility`, `accuracy`) inferred from an explicitly stated justification. For example, a guideline stating "Avoid coloring the axis labels, this can make the data harder to read"[SMR-MARKETING] would be tagged with `clarity`, `simplicity`, and `comprehensibility`.

Rather than selecting a minimal subset, we constructed a *theory-informed value vocabulary* by unifying all values drawn from prior work on Value Sensitive Design [22], visualization ethics [15, 57], visualization guidelines [12], and trust in visualization [36, 42, 56]. We provided the model with a closed set of 34 value categories, detailed in Appendix B, each accompanied by a strict definition, example phrasings, and disambiguation notes (e.g., distinguishing `clarity` from `comprehensibility`, or `accuracy` from `honesty`). Value tagging was restricted to guidelines with explicitly stated rationales (approximately 35% of the corpus). This constraint reduces speculative inference by grounding each tag assignment in the author's stated reasoning. We also minimize cultural or occupational bias in interpreting the less verbose guidelines. For each eligible guideline, the LLM assigned up to three value tags from the 34-value vocabulary that best aligned with the stated rationale.

To assess reliability, we compared LLM-generated tags against a human-coded reference set of 100 guidelines. Of these, 37 contained explicit rationales and were independently tagged by a human annotator using the same 34-value vocabulary and the same "up to three tags per guideline" constraint. Agreement was computed by comparing the sets of value tags assigned to each guideline by the human and the LLM. We report agreement at three levels of strictness:

- **Strict agreement**: exact match between the two tag sets.
- **Mid agreement**: at least two shared tags.
- **Loose agreement**: at least one shared tag.

Under these definitions, agreement was Strict = .216, Mid = .351, and Loose = .919; mean Jaccard overlap was .486 (median .333). Per-tag macro Cohen's $\kappa = .322$ (prevalence-weighted $\kappa = .323$), and pooled micro $F_1 = .616$. Exact matches were relatively uncommon, largely because several values capture closely related concerns (e.g., `clarity` vs. `comprehensibility`). To account for this, we collapsed the 34 values into six higher-level *value families* using a synonym map, analogous to our normalization procedure for action–target pairs. At the family level, agreement improved substantially: exact family-set match = .703, Loose ($\geq$ 1 shared family) = 1, mean Jaccard = .845 (median = 1); pooled micro $F_1 = .678$, macro $\kappa = .325$ (prevalence-weighted $\kappa = .408$).

As with action–target normalization, most disagreements reflected fine-grained distinctions among semantically adjacent concepts rather than fundamentally different interpretations of the guideline rationale. We therefore treat value tags as calibrated heuristic indicators: not definitive ground truth, but sufficiently reliable to support aggregate analyses of value emphasis across organizations, chart types, and domains.

7 Finally, we applied a rule-based layer to the full dataset to further standardize action and target tags and improve reliability. For example, in the guideline "Always start at a zero baseline for the x-axis,"[CMU-EDUCATION] rules validated extract of { *action* : START, *target* : AXIS_ZERO, *strength* : MUST }, enforcing that AXIS and ZERO were combined as AXIS_ZERO. Guidelines sharing the same normalized action–target pairs (i.e., accounting for synonyms and negative vs. positive framing) were then clustered to reveal areas of consensus or variation across organizations.

6 Analysis of Chart-Specific Guidelines

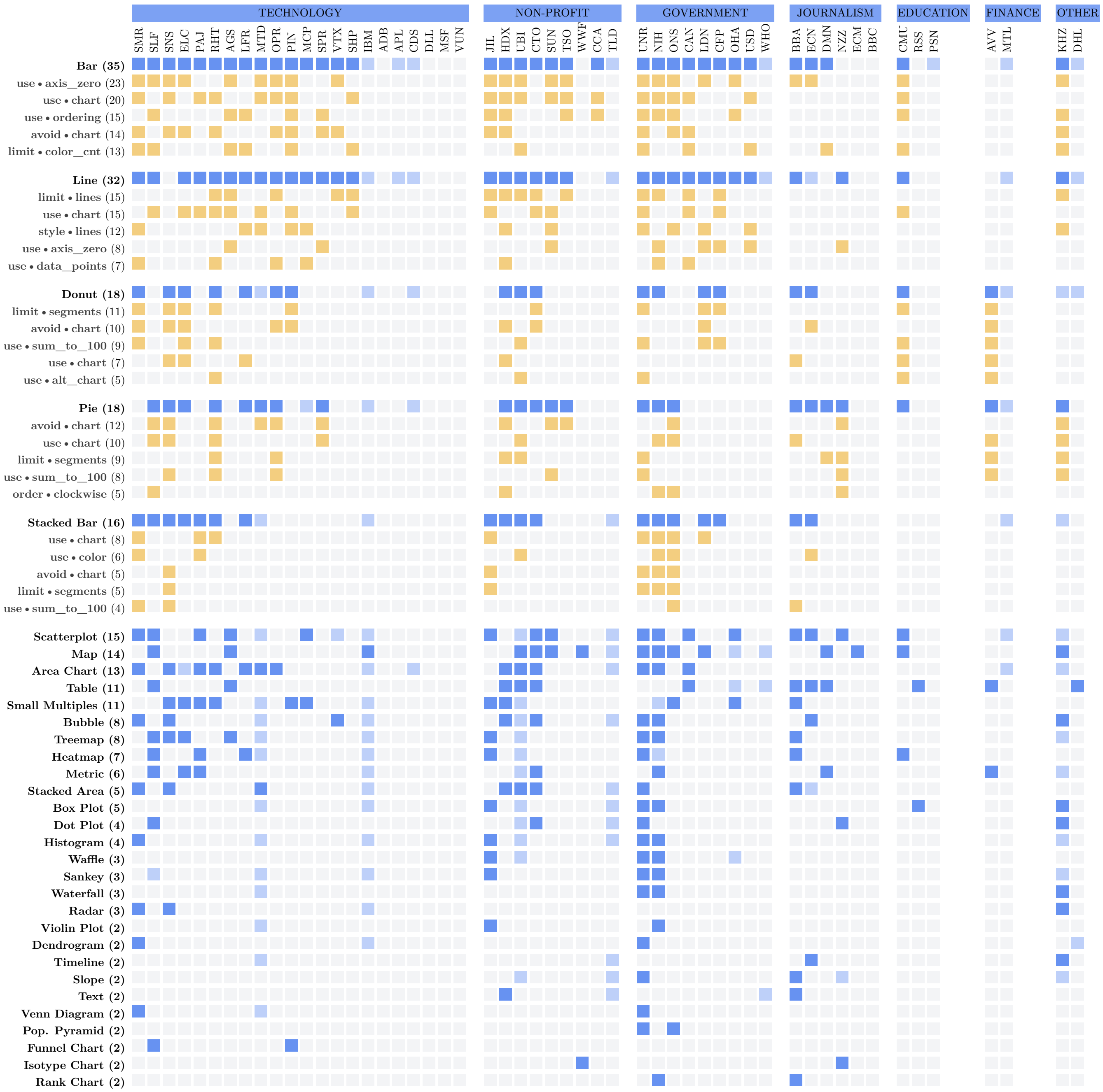
Having examined how organizations structure their style guides and which components they emphasize, we now turn to the content of the chart-specific recommendations themselves. Building on the standardized dataset described in Section 3 and the NLP analysis in Section 5, this section analyzes how guidance varies across chart types and industries. Figure 5 presents an overview in which charts appear as rows and organizations as columns, grouped by industry. For completeness, we also include organizations whose style guides do not provide explicit chart-type sections. As shown in Figure 4, they typically include task descriptions and color guidance.

Among the many chart types covered, four were notably prevalent: *bar*, *line*, *pie*, and *donut* charts. These are not only the most common in everyday practice but also the most contested in design advice, making them a useful lens for comparing consensus and conflict across organizations. As we discuss below, strong agreement emerges for bar charts, consensus fragments for line charts, and pie and donut charts trigger outright disagreement.

6.1 Consensus: Shared Ground Across Industries

Consensus on Bar Charts. A salient pattern in the corpus is the emergence of cross-industry consensus around a small set of design rules for bar charts. Of the 42 style guides with chart-specific recommendations, 35 (83%) included rules for bar charts, making them the most prominent and pervasive chart type. By comparison, line charts were slightly less frequent with 32 (74%), pie charts were mentioned in fewer than half (45%), and scatterplots in only 15 (30%), underscoring the centrality of bar charts in institutional guidance. Style guides typically specify when to use or avoid specific bar chart variants (e.g., vertical versus horizontal layouts), and the majority converge on starting the y-axis at zero, sorting bars in descending order, and limiting the number of colors. Across these bar-chart guidelines, value framings for rationales emphasized `clarity`, `efficiency`, and `accuracy` (88%, 44%, and 37%, respectively), suggesting priorities rooted in readability and faithful magnitude comparison. The frequency and consensus around rules suggest that bar charts are the "safe" default option across diverse organizational contexts.

Figure 5: The top guideline themes. We use ■ when the style guides provide chart guidelines, ■ when it mentions the chart but provides no specific guidelines, and ■ when not mentioned. Charts that appeared in fewer than 2 style guides are excluded. We provide the top guidelines for common charts, with ■ indicating when that specific rule is present.



ADB Adobe Spectrum
AGS Storymaps Arc GIS
APL Apple
AVV AVIVA
BBA BBC Audiences
BBC BBC Global
CAN Government of Canada
CCA Catalyst California
CDS Cloudscape Design System
CFP Consumer Financial Protection Bureau
CMU Carnegie Mellon University
CTO Cato Institute
DHL DHL
DLL Dell Design System
DMN The Dallas Morning News
ECM The Economist
ECN The Economist
ELC Elastic UI
HDX Humanitarian Data Exchange
IBM IBM Design Language
JIL Justice Innovation Lab
KHZ Delish (Kraft Heinz Food)
LDN London City Intelligence
LFR Liferay
MCP MailChimp
MSF Microsoft Office Learn
MTD Material Design 2
MTL MetLife
NIH National Institutes of Health
NZZ NZZ Visuals
OHA US AID | Office of HIV/AIDS
ONS Office for National Statistics
OPR OPattern
PAJ Pajamas Design System
PIN Pinterest
PSN Pearson
RHT RedHat
RSS Royal Statistical Society
SHP Shopify
SLF Salesforce
SMR Semrush
SNS SONOS
SPR Sprout Social
SUN Sunlight Foundation
TLD Trafford Data Lab
TSO TASO
UBI The Urban Institute
UNR United Nations High Commissioner for Refugees
USD U.S. Web Design System
VTX VTEX
VUN Visma's Unified
WHO World Health Organization
WWF World Wide Fund for Nature

Bar Chart

Organizations: 35 **Industries:** 8

Top 7 consensus rules:
- use • axis_zero *(23 orgs, 8 industries)*
- use • chart_type *(20 orgs, 6 industries)*
- use • ordering *(15 orgs, 7 industries)*
- avoid • chart_type *(14 orgs, 6 industries)*
- limit • color_count *(13 orgs, 8 industries)*
- use • color *(11 orgs, 5 industries)*
- use • direct_labels *(9 orgs, 4 industries)*

Convergence on Zero-Baseline for bar chart. Figure 5 shows that the zero-baseline rule is the single most frequent bar chart recommendation in our corpus, appearing in 23 of 35 (66%) guides that included bar chart guidance. In our tags, zero-baseline rationales were predominantly coded as `honesty`, `accuracy`, and `non_maleficence` (e.g., 100%, 100%, and 62%). The dominance of the zero- baseline rule, together with honesty and non- maleficence among the top value indicators, is noteworthy because it reflects convergence between organizational practice and scholarship that cautions against misleading axis truncation. This recommendation is supported by research showing that truncated axes can distort data interpretation [34, 41] and several style guides explicitly reference this concern. The stated rationales for this recommendation include, "starting from a non-zero baseline can distort the ratios between the amounts being compared" [PIN-SOCIAL MEDIA], and "truncating the start of the axis can create deceptively large differences across bars even when the differences are small" [NIH-GOVERNMENT]. At the same time, more recent work suggests that the magnitude of such distortions depends on task and context [35]. However, this analysis suggests that nuance from newer studies is yet to be incorporated.

6.2 Contradictions: Contestation and Divergence

In contrast to the strong consensus on bar charts, guidance on line charts is more fragmented. Of the 32 style guides that mention line charts, many agree on basic practices. We see moderate agreement on recommendations such as limiting the number of lines (15 of 32; 46% of style guides with line chart rules) and styling lines for clear differentiation (9 of 32; 38%), suggesting a shared concern with avoiding "spaghetti plots" and ensuring visual clarity. However, consensus quickly breaks down on the question of zero baselines.

Line Chart

Organizations: 32 **Industries:** 8

Top 7 consensus rules:
- use • chart_type *(15 orgs, 6 industries)*
- limit • lines *(15 orgs, 5 industries)*
- style • lines *(12 orgs, 6 industries)*
- use • axis_zero *(8 orgs, 5 industries)*
- use • direct_labels *(7 orgs, 3 industries)*
- use • data_points *(7 orgs, 3 industries)*
- truncate • axis *(6 orgs, 4 industries)*

While 23 style guides endorse zero baselines for bar charts, only 8 do so for line charts.

Divergence on Zero-Baseline for Line Charts. Six style guides explicitly recommend cutting the axis "to show subtle, yet significant changes that wouldn't be visible when the y-axis is extended to zero" [NZZ-GOVERNMENT]. The same document, however, also warns against manipulative truncation, advising not to "cut the y-axis to make a trend appear more dramatic than it is" [NZZ-GOVERNMENT]. Other organizations propose thresholds, such as "set the y-axis min and max so the line occupies about two-thirds of the range" [AGS-TECHNOLOGY]. Among rationales that endorse axis truncation for line charts, value framings most often emphasized `clarity` (100%) and `comprehensibility` (40%), with no references to `honesty` or `non_maleficence`. In contrast, among the two organizations that included rationales when warning against truncation, both were coded as `honesty`, `accuracy`, and `non_maleficence`. These tensions suggest that the zero-baseline rule is less accepted for line charts, reflecting both limited research attention and potentially differing communicative priorities. Notably, recommendations for zero baselines in line charts disproportionately come from government agencies, aligning with our analysis showing the top institutional value framings as `clarity` (88%) and `accuracy` (42%).

Pie Chart

Organizations: 18 **Industries:** 8

Top 7 consensus rules:
- avoid • chart_type *(12 orgs, 6 industries)*
- use • chart_type *(10 orgs, 7 industries)*
- limit • segments *(9 orgs, 6 industries)*
- use • sum_to_100 *(8 orgs, 6 industries)*
- order • clockwise *(5 orgs, 4 industries)*
- consider • alternative_chart *(5 orgs, 4 industries)*
- use • color *(4 orgs, 4 industries)*

Contention around Pie Charts. Pie charts emerge as one of the most divisive chart types in the corpus. Of the 42 style guides, 18 (45%) include explicit recommendations for pie or donut charts. Where they are addressed, organizations typically treat donut charts as interchangeable with pies, though four distinguish them as a distinct chart type, using the center as a unique feature. A few guides recommend utilizing the empty space to highlight key metrics, e.g., "Use the middle of the doughnut to highlight the number or category type" [CFP-GOVERNMENT].

Among the 18 that included pie or donut charts, 66% caution their use. Some style guides prohibit them outright, stating it "can be hard to compare slices,"[OPR-TECHNOLOGY] while others allow their use with conditions such as limiting slices to 4–5, maintaining consistent ordering, or using them only for simple part-to-whole relationships. Notably, the rationales for avoiding pie charts vary. Some echo academic findings about perceptual difficulty in judging areas, stating "It is much more difficult to visually judge the size of circles (or circle segments) vs. rectangles" [SUN-NONPROFIT]. Others cite pragmatic concerns such as space efficiency, e.g., "It takes a lot of space to show little information…" [HDX-NONPROFIT]. In our value tags, cautionary rationales most frequently invoked `clarity` (100%), `efficiency` (94%), and `accuracy` (63%). In contrast, permissive-but-conditional guidance invoked `clarity` (100%) alongside `simplicity` (71.43%) and `storytelling` (29%), e.g., "Use to turn very simple data points (59% says 'yes') into a compelling visual"[SPR-TECHNOLOGY] or to "call out a particularly important point" [BBA- JOURNALISM]. Thus, pie charts exemplify how style guides encode not only perceptual considerations grounded in empirical research that prioritize speed and accuracy measures [14, 26, 52], but also contextual and pragmatic factors that reflect the priorities of different organizations.

Diversity in Chart Types Covered. Beyond canonical forms like bars, lines, and pies, style guides also address a long tail of specialized charts, including *Sankey diagrams*, *radar charts*, *waterfall plots*, and *violin plots*. Their inclusion suggests that some organizations are motivated to codify best practices even for rare or advanced visualization types. Further inquiry is needed to disambiguate whether that wide variety of charts reflects observed internal usage, aspirational coverage, or emerging visualization trends.

While chart-specific guidance highlights localized areas of agreement and disagreement, this analysis leaves open central questions that motivated this project, i.e., *how consistent are guidelines across organizations and industries, and what broader norms do they embed?* To address this, we examine guidelines at a chart-agnostic level by clustering them into shared action–target pairs, thereby enabling us to identify broader design norms and contradictions rooted in organizational values.

7 Analysis of Chart-Agnostic Guidelines

Our analysis uncovered a set of chart-agnostic heuristics that apply across visualization types. We define a guideline as *chart-agnostic* if it is recommended for at least five distinct chart types. These chart-agnostic guidelines address practices like labeling, color usage, and data simplification. With the exception of the zero-baseline recommendation, they are generally less debated and highly pervasive, potentially revealing the underlying norms that organizations view as general principles for visualization design. We identify four main themes based on the top chart-agnostic guidelines clusters in our corpus (Figure 6): color, simplicity, accuracy, and readability.

7.1 Color as Necessity and Risk

Color emerges as one of the most frequently discussed topics across style guides, simultaneously positioned as indispensable and potentially problematic. Color is not treated as decoration alone. Many guides recommend using color to draw people's attention to a particular message (**USE•COLOR**) but warn against overuse, frequently stressing the need to limit the number of colors (**LIMIT•COLOR_COUNT**). This focus likely reflects the inherent complexity of color as an aesthetic, perceptual, and communicative tool.

Most style guides (30 out of 53) include explicit color palettes, often tied to organizational branding or specific data types (e.g., categorical, sequential, divergent, etc.), which shape their recommendations. Accessibility is another recurring concern. Style guides often recommend high-contrast and colorblind-friendly palettes, emphasizing that color choices should support diverse audiences.

Glossary of chart-agnostic guideline clusters.

Theme: Color
use • color: How and when to apply color.
limit • color_count: Limit the number of colors.

Theme: Simplicity
limit • category_count: Limit the number categories.
limit • segments: Limit the number of segments.

Theme: Accuracy
use • axis_zero: Start numerical axes at 0.
use • axis: Recommended use of axes and tick marks.
ensure • sum_to_100: Percentages should sum to 100%.

Theme: Readability
use • direct_labels: Place labels on or near data points.
use • legend: How and when to use legends.
use • ordering: Arrange data logically and intentionally.

Cluster	Orgs	Ind.	GLs	% Orgs
limit • categories	12	4	25	23%
limit • color_count	25	9	44	47%
limit • segments	19	8	34	36%
use • axis	12	6	27	23%
use • axis_zero	26	7	72	49%
use • color	28	8	85	53%
use • direct_labels	18	4	50	34%
use • legend	13	5	35	25%
use • ordering	16	6	32	30%
use • sum_to_100	19	7	38	36%

Notes: Orgs = distinct organizations; Ind. = industries; GLs = guideline instances. % Orgs = share of all 53 orgs with ≥ 1 rule in the cluster.

For example, one government guide emphasizes designing "to differentiate variables for audiences with visual impairments" [USD-GOVERNMENT].

Underlying these cautions is a shared belief that color should be used with intention (about 88% of stated rationales mention a purpose) and in moderation, evidenced by **LIMIT•COLOR_COUNT** ranking among the most frequent guidelines. For instance, one social media organization states, "Keep it to 4 colors maximum, otherwise charts will get hard to read" [PIN-SOCIAL MEDIA]. In bar charts, the most common color-related recommendation is to use only one. In general, color is a workhorse with three vital roles:

(1) **Cohesion and brand reinforcement**. Nearly all style guides provide explicit palette specifications, and 70% offer substantial guidance. In this role, color is used to create a standardized, cohesive visual identity and to ensure consistency across outputs. Some style guides also include explicit rules for aligning visualizations with organizational identity, e.g., ("…use CMU RED to represent CMU" [CMU-EDUCATION]).
(2) **Analysis focus**. About 45% of color-use rationales relate color to analytic tasks. Here, color is used to support comparisons, emphasize patterns, or group related categories within or across plots. For example, "use contrasting or meaningful

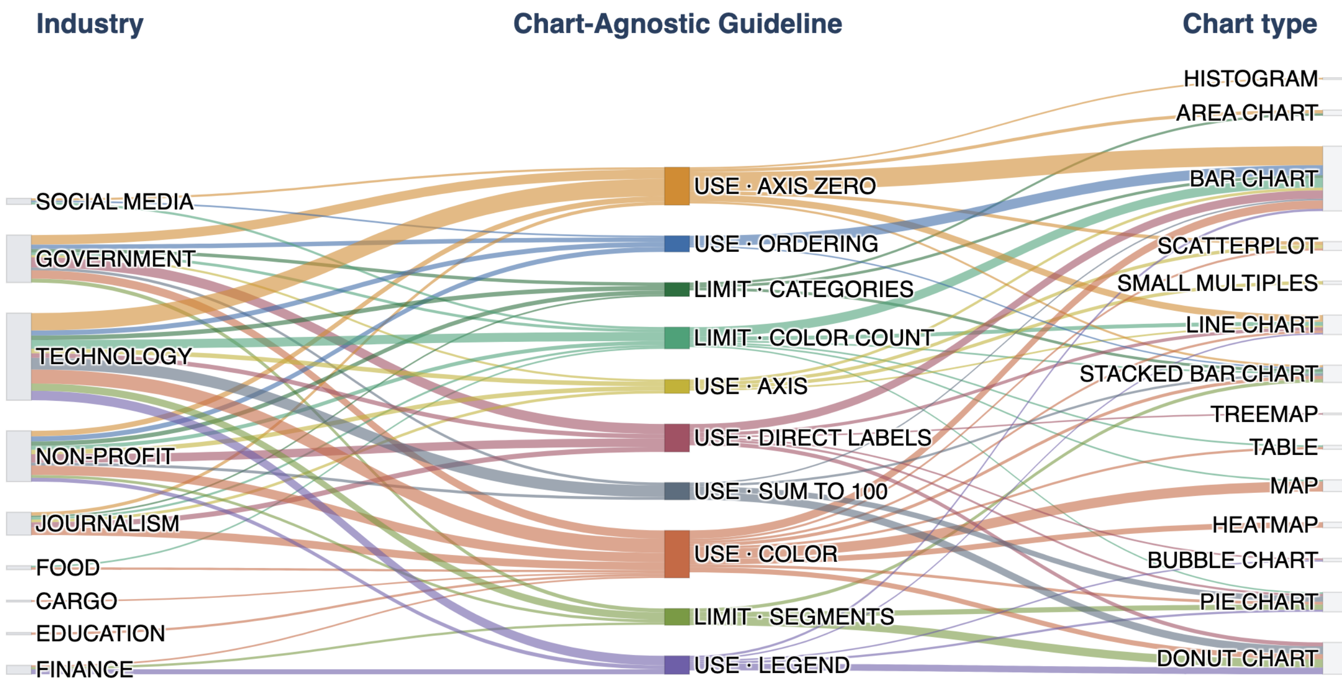


Figure 6: The Sankey diagram illustrates the *chart-agnostic* guidelines in the center, the industries (on the left) that our corpus associates with these design recommendations, and the chart types (on the right) to which the guidelines apply. We define a guideline as *chart-agnostic* if it is recommended for at least five distinct chart types. In addition to the zero baseline rule (USE•AXIS_ZERO), there is agreement on the use of color (USE•COLOR and LIMIT•COLOR) and the recommendation to directly label the charts (USE•DIRECT_LABELS).

colors to show differences in data" [SLF-TECHNOLOGY] and "If categories form meaningful groups, use colour to reinforce that" [ECN-JOURNALISM].

(3) **Emphasis and informative use**. Roughly a third of rationales (29%) emphasize color's role in guiding attention. Color is used to highlight key messages, call out outliers, or mark points of interest. Examples include "Start with grayscale and use color to highlight important points" [CCA-NONPROFIT] and "Use color to highlight points of interest on your map." [NIH-GOVERNMENT].

Taken together, these recommendations reflect a broad consensus: color is too powerful a tool to be used casually. Its role in emphasis, coherence, and branding makes it central to effective communication, and designers are tasked with applying it sparingly to avoid confusion while maximizing accessibility.

7.2 Simplicity as a Guiding Principle

A second recurring theme centers on restraint or keeping charts simple and focused. Many guidelines stress limits on categories and segments (**LIMIT•CATEGORY_COUNT**, **LIMIT•SEGMENTS**), particularly for categorical data. Related rules such as **LIMIT•LINES** for line charts and **LIMIT•DATA_POINTS** for scatterplots or bubble charts reflect the same principle of avoiding overwhelming readers with visual clutter. In our dataset, rationales for these limits universally cite clarity (100%), simplicity (100%), and efficiency (89%).

Across chart types, organizations treat simplicity as a safeguard against overload. This principle is operationalized numerically, recommending limits on pie slices, line series, or scatterplot points. However, the exact thresholds vary widely, from as few as 2 categories or lines to as many as 7. A common fallback strategy is to "Aggregate smaller categories into an 'other' or 'miscellaneous' category" [UBI-NONPROFIT]. Although this principle may be most relevant for communicative and public-facing contexts, consensus around this topic reflects a common belief that effective visualization requires restraint, prioritizing clarity and minimalism over exhaustiveness.

7.3 Accuracy and Guardrails Against Design Hazards

A third cluster of rules emphasizes accuracy and integrity. Similar to our *chart-specific* analysis in Section 7, one of the most frequent *chart-agnostic* recommendations is the use of zero baselines (**USE•AXIS_ZERO**) to prevent exaggeration of small differences. Related guidelines include proper axis construction (**USE•AXIS**) and ensuring that percentages sum to 100% when displayed (**ENSURE•SUM_TO_100**). In our coding, clarity (71%), accuracy (69%), and honesty (38%) are the most frequent value tags attached to axis- and baseline-related rationales. Together, these principles underscore a shared organizational commitment to preventing misleading, incomplete, or inaccurate representations.

From Figure 6, we can observe that the zero-baseline rule especially appears in guidelines for *line charts*, *area charts*, *histograms*, *scatterplots*, and *stacked bar charts*. Rationales for these rules consistently reference `accuracy` (95%), `honesty` (85%), and `non_ maleficence` (55%). While its empirical grounding is strong for ratio-related tasks with bar charts (as discussed in Section 6), its extension beyond bar charts suggests that many organizations adopt a risk-averse stance, preferring caution even in contexts where the empirical basis is less clear. Axis-related rules often highlight tick values and consistency, such as "Use natural increments on the Y-axis that are divisible by 1, 2, or 5" [HDX-NONPROFIT], or cross-chart comparability: "Use the same axis units, axis ranges, and chart dimensions across charts in a small-multiples group to facilitate comparison" [JIL-NONPROFIT]. Similarly, the guideline that percentages sum to 100% reinforces the expectation that part-to-whole charts must fully account for the data, thereby minimizing misunderstandings.

These principles reflect the strongest alignment with academic literature. Axis truncation, inconsistent tick marks, and incomplete part-to-whole depictions are canonical examples in research on deceptive or misleading charts [24, 33, 34, 37, 41]. Their prominence in style guides and the value framings attached to them suggests that some academic "design hazards" have successfully translated into industry norms.

7.4 Readability and Effortless Interpretation

Finally, many chart-agnostic recommendations converge on the principle of readability. Common rules include using direct labels where possible (**USE•DIRECT_LABELS**), presenting legends appropriately (**USE•LEGEND**), and ordering categories logically (**USE•ORDERING**). Together, these practices aim to reduce cognitive effort, ensuring readers can interpret charts without unnecessary scanning or guesswork. Across these chart-agnostic guidelines with rationales, readability-related values such as clarity (100%), efficiency (70%), and comprehensibility (45%) were the most frequently invoked tags. This cluster shows a strong convergence toward reader-centered values.

Organizations consistently emphasize that direct labeling is preferable. Some frame this as a matter of precision: "If exact numbers are important, the bars should be labeled with the value" [CFP-GOVERNMENT]. Others stress efficiency or simplicity, noting that "Direct labelling of values eliminates the need for grid lines" [HDX-NONPROFIT]. Similarly, style guides discourage reliance on legends when labels can be placed directly on the chart: "Label groups directly to avoid using a separate legend when possible" [JIL-NONPROFIT]. In our tagging, rationales for direct labeling almost universally invoked clarity (100%) and frequently efficiency (91%), underscoring its role in reducing perceptual overhead.

Ordering is also treated as essential for guiding interpretation and avoiding arbitrary sequencing. Rationales for ordering categories likewise emphasized clarity (100%), efficiency (81%), and comprehensibility (61%), reflecting a shared belief that ordering may be a cognitive scaffold rather than a stylistic choice. The most common advice is to order categories by magnitude. For example, "Sort data from highest to lowest to enable quick comparison" [AGS-TECHNOLOGY] or "The ordering of the bars makes the data easier to understand" [CCA-NONPROFIT]. Exceptions are noted when a natural order exists, such as chronological or categorical conventions.

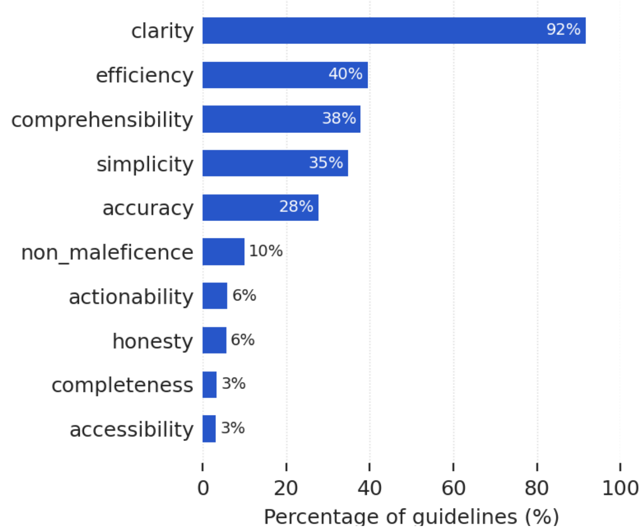


Figure 7: Top 10 value tags across all guidelines with rationales.

Synthesis and Implications

Across all guidelines and chart types, the value framings provide the clearest corpus-level indication of what organizations deem most important in visualization practice. As summarized in Figure 7, `clarity` is the dominant value (92%), followed by `efficiency`, `comprehensibility`, and `simplicity` (40–35%). Ethical or protective values such as `accuracy` and `non_ maleficence` surface less often but are particularly prominent in rules that aim to prevent misleading designs. The implication is that, as they are written today, style guides function as instruments of reader-centered usability. They emphasize making visualizations easy to read, quick to parse, and visually manageable. For practitioners and researchers, these patterns highlight the need to interrogate whether a focus on clarity alone is sufficient for trustworthy communication, and where additional structures are necessary to ensure that usability does not override accuracy, fairness, or transparency.

Although the analysis above provides a bird's-eye view of consensus and contradiction, it inevitably compresses nuance. It obscures the fact that many recommendations are conditional, tied to exceptions, or phrased with varying thresholds. To support transparency and enable further inquiry, we developed the *Guidelines Explorer* as a secondary contribution. The tool is not the focus of this paper, but it serves as a practical interface to the dataset, allowing researchers and practitioners to interrogate the underlying guidelines, compare organizational approaches, and explore context-specific details that go beyond the scope of our aggregated analysis.

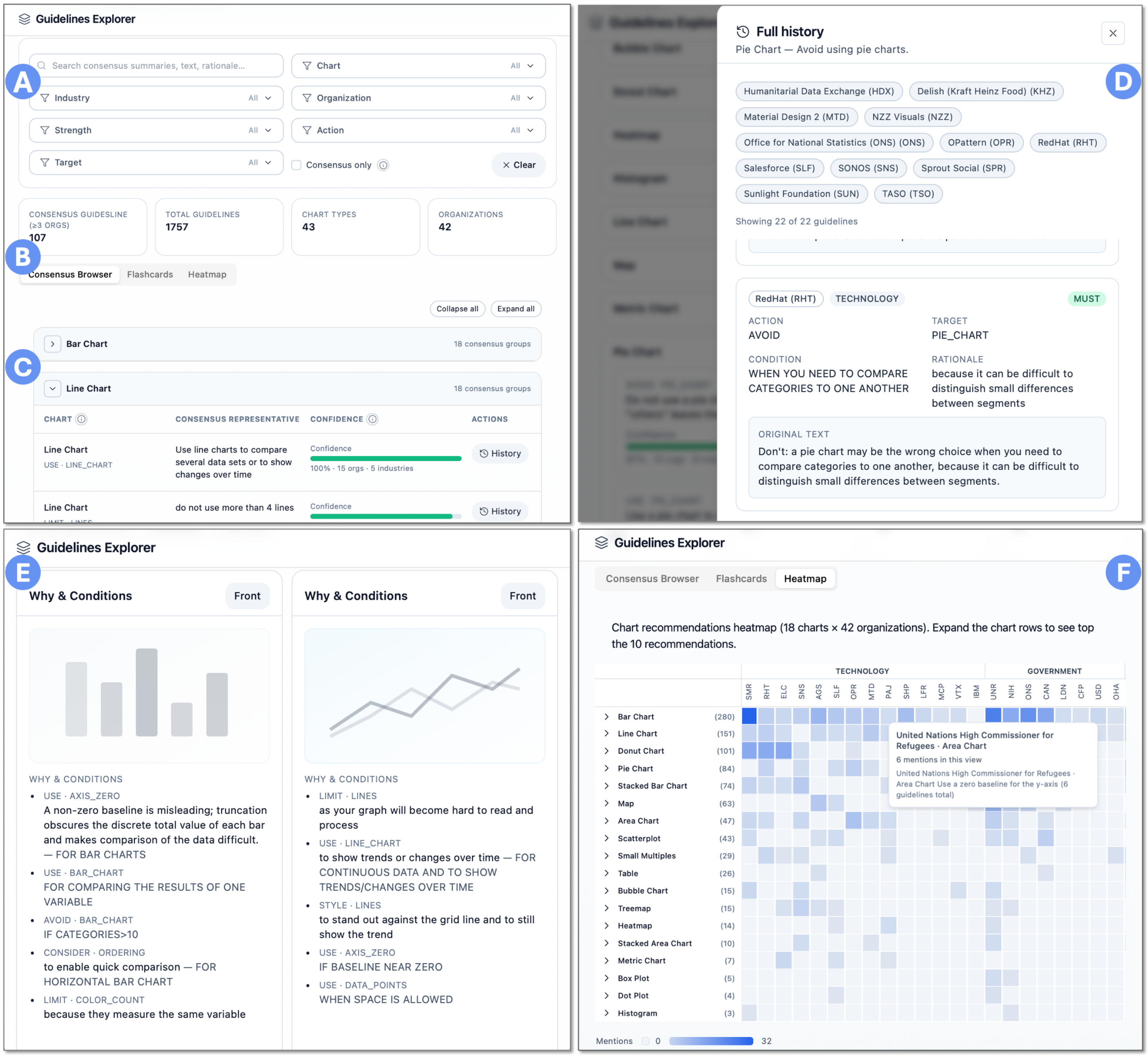


Figure 8: The *Guidelines Explorer*, an interactive tool developed to accompany our analysis. A The interface allows users to filter guidelines by chart type, industry, organization, action, or strength. B Tabs allow users to choose between exploring consensus groups and individual guidelines, viewing flashcard summaries for each chart, or an overview heatmap. C The *Consensus Browser Tab* presents chart-specific consensus groups, it shows the individual guidelines for a specific cluster, and provides a confidence estimate based on the unique industry and organization coverage in the corpus. D Clicking a history button shows a slide-out panel that lists all guidelines associated with that cluster. The *Flashcard Tab* summarizes the top recommendations for each chart. E shows the backs of the cards shown in Figure 1, which details the rationale and conditions attached to each recommendation. F The *Matrix Tab* shows the distribution of chart mentions across industries and organizations. Together, these provide an open and extensible platform for navigating the diversity of visualization guidelines.

8 Guidelines Explorer

The *Guidelines Explorer*[4] was developed to make our corpus of visualization guidelines accessible and actionable for diverse audiences. Although the dataset provides a foundation for comparative analysis, a static table or appendix would limit its usefulness. Our goal was to create an interface that supports exploration, teaching, and practice. The interface is informed by three primary user considerations:

- Researchers (U1): Researchers studying visualization practice need tools to query the dataset and analyze organizational differences.
- Practitioners and Novice Designers (U2): Practitioners and novice designers in finding relevant recommendations.
- Educators (U3): Instructors teaching data visualization who may need to highlight both best practices and contested debates.

8.1 Design Features and Views

The interface was implemented in React with *TypeScript* for type safety and modularity. We used *Tailwind CSS* for styling and responsive layout design, and integrated *shadcn/ui* components for consistent interaction elements (e.g., cards, dropdowns, search bars). The site is deployed using GitHub, providing open access to allow future research and development.

There are three complementary views (Figure 8, Labels C, E, and F), each tailored to support different user tasks while drawing on the same shared corpus. We connect these views to user goals (U1–U3) to show how the design enables researchers, practitioners, and educators to meet distinct needs within a shared environment.

Search and Filter. (A | U1, U2, & U3). This panel serves as the entry point to the corpus. Users can filter by chart type, industry, organization, guideline strength, action, and target, either singly or in combination. Filters propagate across all other views to enable a seamless exploration workflow. For example, a user may filter for "pie charts" and immediately see consensus clusters, flashcards, and heatmap coverage update in concert. Query fields support both structured selection (checkboxes and dropdowns) and free-text search. We designed this dual interaction to accommodate different information-seeking behaviors: researchers conducting systematic reviews (U1) may prefer structured filtering, whereas practitioners exploring design questions (U2) may rely on keyword search.

Consensus Browser. (C | U1, U2, & U3). This view highlights *consensus clusters*, and users can toggle to reveal non-clustered guidelines. Clusters are formed through fuzzy mapping to resolve lexical variation. For example, USE, CONSIDER, and SET are aliased as USE, while COLOR and COLOR_PALETTE are unified as COLOR. A cluster qualifies as consensus when it includes at least three guidelines from three distinct organizations. This threshold is chosen to ensure that clusters represent shared practice rather than idiosyncratic style. Each cluster can be expanded to reveal its constituent guidelines, organizations, and verbatim texts as illustrated with D.

To further support novices (U3), clusters include a confidence indicator. Full confidence is defined as reaching 70% organization coverage and 30% industry coverage for style guides with the relevant chart. The inclusion of this indicator aims to encourage users to interpret clusters not as rigid rules but as varying in support. For example, a researcher might use the consensus view to identify clusters strong enough to warrant a comparative study (U1), while an educator may use the confidence indicator to illustrate that even widely cited rules are context-dependent (U3).

Flashcards. (E | U2 & U3). Each card summarizes the top five consensus clusters for a given chart type. The card front lists the guidelines, and the back includes rationale, common exceptions, and alternative practices where available. This flashcard metaphor was chosen for its familiarity in pedagogical contexts and its affordance to distill the complex data into digestible insights. Designers under time pressure (U2) may quickly consult the top recommendations for a chosen chart type, while educators (U3) may use "conditions" and "exceptions," to reinforce that guidelines are context-dependent.

Heatmap Overview. (F | U1 & U3). This aggregate view presents guideline distributions across organizations and chart types in a matrix format. Each cell is shaded by the number of guidelines, allowing patterns of dominance and divergence to emerge at a glance. Cells are interactive, and hovering shows example guidelines that are relevant. The heatmap enables educators (U3) to illustrate industry-level patterns, such as the heavy use of direct-labeling guidelines in government compared to their rarity in technology, while researchers (U1) can identify underrepresented chart types or industries as potential gaps in the literature.

8.2 Implications

Considering the usability of our corpus, the volume, diversity, and uneven coverage of guidelines make it difficult to understand how individual rules relate to broader patterns. A single recommendation from a single organization is rarely generalizable, and meaningful insights emerge only when guidelines are viewed collectively, compared across contexts, and situated within larger patterns of consensus or contradiction. The *Guidelines Explorer* addresses this challenge by enabling:

- Synthesis of dispersed guidance. Users can surface areas of agreement and disagreement across dozens of organizations rather than relying on isolated recommendations.
- Cross-organizational and cross-industry comparison. Filters for chart type, industry, and organization expose how similar rules manifest differently across domains, revealing which norms are widely shared and which are domain-specific.
- Inspection of rationale and value framings. Each guideline links directly to its justification and associated values, helping users understand why a rule exists, not just what it prescribes.
- Exploration of chart-agnostic clusters. Grouping guidelines (e.g., ordering, baseline practices, color limits) illuminates recurring design principles that cut across chart types and industries.
- Pedagogical and interpretive support. Summary views, flashcards, and cluster inspections provide approachable ways to illustrate how visualization norms emerge, evolve, and diverge.

[4] https://washuvis.github.io/styleguides/

In summary, the Guidelines Explorer complements our quantitative analysis by making the underlying corpus navigable and interpretable. It enables users to compare organizational practices, trace the provenance of specific consensus recommendations, inspect value framings, and explore chart-agnostic patterns that would otherwise remain hidden. While not the focus of this paper, the tool provides an essential bridge between aggregated findings and the nuanced, context-dependent nature of real-world visualization guidance.

9 Discussion

Visualization style guides are often treated as straightforward repositories of "best practices," i.e., technical rules distilled from perceptual research and refined through experience. Our analysis complicates this view. Across 53 organizations, we show that guidelines are not merely instructional artifacts but socio-technical documents that encode institutional priorities, cultural norms, audience assumptions, and risk considerations.

9.1 There is no single template, but there are core expectations.

Structurally, the guides in our corpus differ substantially, both within and across industries. Although this heterogeneity suggests that there is no single blueprint for what a style guide should contain, we observe several components that appear in a majority of guides: charts, color, tasks-per-chart, and general visualization guidelines. This supports and extends prior work by Kandogan and Lee [30], who found, in their analysis of academic and book-derived guidelines, that design rules often extend beyond chart types and encodings. Our findings similarly point to a broad and holistic view of visualization practice.

One interpretation of this structural variation is that style guides offer a window into the differing priorities of industries, reflected in what they emphasize or omit. For example, substantial coverage of tasks may indicate a stronger analytics focus, as we observed in many *Government* guides. In contrast, richer guidance on annotations, color, and typography may signal a storytelling orientation, as seen in Journalism organizations such as the US-based *The Economist* and Switzerland's *Neue Zürcher Zeitung*. At the same time, we also observe high variance within industries, which likely reflects the diversity of organizations themselves. *Technology*, for instance, spans Social Media (e.g., Sprout Social), E-Commerce (e.g., Shopify), and hardware companies (e.g., SONOS). This variation highlights a recurrent challenge in our analysis, i.e., balancing granularity and abstraction when comparing heterogeneous documents, which we will discuss more in Section 9.3.

An alternative, equally plausible explanation for the observed heterogeneity is that it reflects differing levels of investment in documentation and design governance. Even if we assume limited or uneven investment, the presence and consistency of certain core components still provide meaningful insight. For instance, even organizations with limited design governance consistently include charts and color guidance. This suggests that even when developing minimal-effort guides, they should include stable rules for these foundational decisions, which have high downstream impact on quality and consistency. However, researchers must be careful when generalizing findings or proposing universal standards. Our analysis suggests that writing a style guide is not merely a documentation task but a form of organizational self-definition. Explicitly choosing what to include and what to omit offers an opportunity to align the guide with institutional goals and the needs of its audiences.

9.2 Guidelines encode priorities, assumptions, and trade-offs.

In our analysis of individual guidelines, we take a different approach from prior work, which has investigated the prescriptive nature of guidelines (i.e., the use of "dos and don'ts") [18] and the structural properties of guidelines that could support automatic generation [12]. Instead, we conduct a systematic, cross-industry analysis of agreements and conflicts, motivated by the goal of creating a repository of standard visualization guidelines that helps users understand not only the recommendations themselves but also their context, rationales, exceptions, and other sources that might support or contradict the guidance. This need is especially acute given prior work showing that visualization guidance is highly opinionated and often fails to explicitly cite empirical evidence [17]. Our analysis of how organizations articulate, justify, and operationalize visualization rules reveals the following insights:

Broad consensus exists around a small, stable set of norms. Despite wide variation in structure and emphasis, certain rules recur across industries. For example, 66% of guides with bar-chart recommendations mandate a zero baseline, and most industries converge on limits for color count, category count, and the use of direct labeling over legends. These high-frequency rules form a core set of shared design norms that emerge across government, journalism, technology, nonprofit, and corporate guides, suggesting they form a widely accepted foundation for effective visualization. Organizations creating new guides can confidently include these as baseline recommendations.

Guidelines are value-shaped, not value-neutral. Across all guidelines with explicit rationales, clarity appears in 92%, efficiency in 40%, comprehensibility in 38%, and simplicity in 35% of value tags. Government agencies additionally foreground clarity (89%), accuracy (42%), comprehensibility(37%), whereas journalism emphasizes clarity (87%), comprehensibility (38%), simplicity (36%), and the top technology value signals were clarity (91%), comprehensibility (40%), efficiency (38%). These patterns empirically demonstrate that visualization rules encode institutional priorities as much as perceptual considerations. Therefore, new style guide authors should assume that every rule they write communicates why a particular design choice matters. Making those values explicit can help prevent misinterpretation, e.g., distinguishing when simplicity is intended to reduce cognitive load versus when accuracy must take priority.

Contradictions arise from competing institutional priorities. We observed numerous cases where organizations prescribe opposite behaviors, for instance, permitting multiple pie charts for comparison [AVV-FINANCE] versus prohibiting them outright [HDX-NONPROFIT]. Our value-tag analysis shows that permissive guidelines tend to emphasize storytelling and simplicity, whereas restrictive guidelines foreground accuracy and efficiency. These conflicts reflect value trade-offs, suggesting different organizations optimize for different communicative risks and goals. Style guide authors should therefore articulate when and why a recommendation applies and avoid presenting opinion as a universal principle. Doing so can reduce reader confusion and help organizations avoid unintentionally misleading or overly rigid prescriptions.

While these findings illuminate how organizations converge and diverge in the rules they prescribe, the underlying data analysis required making several methodological choices, such as restricting the analysis to guidelines with explicit rationales, defining a controlled vocabulary of values, and using a hybrid human–LLM coding process that balanced interpretive nuance with scalability, discussed next.

9.3 Scaling Qualitative Analysis with Constrained LLMs

Prior work on visualization guidelines often uses qualitative approaches such as grounded theory, which work well for small, focused corpora [12, 18]. However, our corpus comprises 53 organizations and approximately 2,120 individual guidelines across various industries, geographies, and documentation traditions. The scale and heterogeneity of this dataset introduced methodological challenges that motivated a different approach. We adopted a hybrid human–LLM pipeline: *spaCy* enumerated the linguistic space of actions and targets, a human researcher refined and collapsed clusters, and GPT-5 applied a controlled vocabulary of values with deterministic settings. This division of labor reflects our choice to use NLP tools for scaling while tightly constraining inference. For example, we set the model temperature to 0 to ensure deterministic output and provided exact definitions and closed tag sets. In this way, the human analyst defined and constrained the conceptual space within which automated structure was applied. This approach, i.e., constraining LLM inference with deterministic parameters and controlled vocabularies, can provide a blueprint for reproducible, large-scale coding. This may help visualization researchers scale qualitative analyses traditionally limited by time, labor, and coder training.

For values, we curated a theory-informed controlled vocabulary of 34 terms. To the best of our knowledge, there is no existing set of visualization-specific values. Our approach was to draw on the body of prior work on value-sensitive design [22], rhetorical visualization [29], ethical visualization [15], and trust [36]. Even when some values did not appear immediately relevant to visualization (e.g., solidarity, defined as supporting social cohesion, collective welfare, or community representation), we retained these terms for two reasons. First, scholars in value-sensitive design caution against assuming fixed or universal value sets and warn of cultural biases in value interpretation. Second, casting a wide net allows absence to be informative, as absent values could signal what organizations do not emphasize. To that end, several values often associated with data communication, such as transparency, trust, and objectivity, were rare in our corpus, despite their prominence in academic discourse. Further inquiry is needed to disambiguate this phenomenon.

Still, the need to construct our own value vocabulary highlights a gap in the visualization literature, as there is no shared theoretical account of the values that shape design recommendations. Our findings suggest an opportunity for researchers to develop more formal frameworks of visualization-relevant values, grounded in empirical data about practice rather than abstract assumptions. Additionally, values are inherently interpretive constructs, which further complicates matters when organizations do not document why a recommendation exists and under what conditions it may change. Even with explicit rationales, authors often gesture toward multiple priorities at once, and organizational or cultural norms may shape phrasing in subtle ways. We therefore encourage readers to treat value tags as heuristic indicators of the concerns that organizations foreground. These are useful for comparative analysis but not exhaustive representations of their internal design philosophies. Overall, this methodology enabled us to identify large-scale patterns without sacrificing conceptual rigor, providing a model for future research on practice-derived corpora.

9.4 Alignment and Divergence Between Research and Practice

Our findings align with Hullman's argument that visualizations are rhetorical devices [29]. The implication is that empirical visualization research must engage not only with perceptual accuracy but also with *values, audiences, workflows, and organizational risk profiles*. For visualization research to be applicable in practice, it is imperative to not just ask "what makes an effective chart?" but we should interrogate "what chart is effective *for whom*, under what constraints, and in service of what values?" Indeed, this aligns with long-standing calls and ongoing research efforts to acknowledge contexts and individual differences explicitly rather than assume universality.

Certain academic findings have clearly diffused into organizational practice. Zero baselines for bar charts, colorblind-safe palettes, ordering categories for comparison, or chart-type selection, and limits on category counts all appear frequently and consistently across industries, particularly in government and nonprofit guides. These areas reflect strong and longstanding empirical consensus [14, 26, 52]. Similarly, skepticism about pie charts is present in both communities. However, there are notable divergences. For example, recent academic results emphasize task dependency and magnitude sensitivity when considering the zero baseline rule [35], but many organizations often prescribe zero baselines universally. Research on color often focuses on perceptual factors, but in practice, color use may be governed by brand and narrative priorities. Future research might consider color selection under constraints such as brand coherence and visual identity.

Other practitioner concerns reveal empirical gaps ripe for research. For example, style guides frequently specify numeric thresholds such as "no more than four line series," "limit to 4–5 pie slices," "restrict scatterplots to avoid overplotting." However, these thresholds vary widely and rarely cite evidence. These inconsistencies reveal areas where academic research has limited or conflicting results, especially around category limits, line clutter, mixed baselines, and emphasis techniques. The guides thus highlight research questions that matter in practice but are understudied academically. These divergences reveal that real-world visualization must satisfy broader communicative goals, including organizational, narrative, ethical, and aesthetic ones, which academic guidelines often underemphasize.

10 Limitations and Future Work

While our analysis provides a unique view into organizational practices and *Guidelines Explorer* provides a novel platform for exploring these practices, it has several limitations that inform directions for future work.

Scope of Corpus. Our dataset includes 53 publicly accessible style guides, but many organizations maintain internal or proprietary guidelines that we could not access. As a result, our corpus is not exhaustive and may underrepresent industries such as healthcare or finance. Expanding the dataset to include these sectors, as well as guides published in languages other than English, is an important next step. Additionally, the analysis in the paper only scratches the surface and excludes important considerations, such as accessibility, color, interactive, and responsive design, presenting a fertile ground for future work.

Standardization and Intent Extraction. Although we used a mixed human-AI approach to intent extraction, this process is imperfect. Subtle differences in tone or context may have been flattened in the process of converting natural language into action–target units. Moreover, there are numerous instances where multiple actions and targets exist in a single sentence. Future work could explore more nuanced natural language processing techniques or crowdsourced validation to capture these subtleties.

Lack of Formal Evaluation. While we designed the *Guidelines Explorer* interface with researchers, practitioners, and educators in mind, we have not yet conducted a systematic evaluation of how these groups actually use the tool. A careful evaluation would involve formative studies with educators in classroom settings, practitioners in applied contexts, and researchers conducting meta-analyses of design guidelines. Such evaluations would help refine the interface and assess its real-world value.

Evolving Guidelines. Visualization practices are not static; they evolve alongside new technologies, cultural shifts, and emergent data practices. The analysis in this paper and *Guidelines Explorer* presents a snapshot of organizational guidance, but future work could extend it into a living corpus that tracks how guidelines change over time and highlights temporal trends in consensus and conflict.

11 Conclusion

The work in this paper demonstrates that style guides are more than collections of tips. Instead, they are cultural artifacts that both reflect and shape the visualization ecosystem. By aggregating guides across domains, we uncovered shared conventions that function as de facto norms, revealed contradictions that encode sector-specific values, and traced the influence of research as it filters into professional practice. These patterns would not have been visible without treating style guides as a collective phenomenon or without the companion web-based *Guidelines Explorer*. Beyond documenting what guidelines say, the process of comparison itself generated new knowledge, demonstrating how visualization advice circulates, stabilizes, and diverges across communities.

Acknowledgments

This work was partially supported by the National Science Foundation (NSF) under Award Nos. 2142977 and 2330245, which support the Engineering Research Center for Carbon Utilization Redesign through Biomanufacturing-Empowered Decarbonization (CURB).

References

[1] Esha Adhya. 2015. Key elements of an effective style guide in the new age. *Technical Communication* 62, 3 (2015), 183–192.
[2] Muna Alebri, Enrico Costanza, Georgia Panagiotidou, and Duncan P Brumby. 2023. Embellishments revisited: Perceptions of embellished visualisations through the viewer's lens. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2023), 1424–1434.
[3] Sara Alspaugh, Nava Zokaei, Andrea Liu, Cindy Jin, and Marti A Hearst. 2018. Futzing and moseying: Interviews with professional data analysts on exploration practices. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 22–31.
[4] A Arunkumar, L Padilla, and C Bryan. 2025. Modeling and Measuring the Chart Communication Recall Process. In *Computer Graphics Forum*. Wiley.
[5] Jacques Bertin. 1967. Sémiologie graphique: les diagrammes, les réseaux, les cartes. *(No Title)* (1967).
[6] Alex Bigelow, Steven Drucker, Danyel Fisher, and Miriah Meyer. 2014. Reflections on how designers design with data. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*. 17–24.
[7] Rita Borgo, Alfie Abdul-Rahman, Farhan Mohamed, Philip W Grant, Irene Reppa, Luciano Floridi, and Min Chen. 2012. An empirical study on using visual embellishments in visualization. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2759–2768.
[8] Willard Cope Brinton. 1919. *Graphic methods for presenting facts*. Engineering magazine company.
[9] Stuart Card. 2009. Information visualization. In *Human-computer interaction*. CRC press, 199–234.
[10] Bodong Chen and Haiyi Zhu. 2019. Towards value-sensitive learning analytics design. In *Proceedings of the 9th international conference on learning analytics & knowledge*. 343–352.
[11] Qing Chen, Ziyan Liu, Chengwei Wang, Xingyu Lan, Ying Chen, Siming Chen, and Nan Cao. 2022. Vizbelle: A design space of embellishments for data visualization. *arXiv preprint arXiv:2209.03642* (2022).
[12] Jinhan Choi, Changhoon Oh, Bongwon Suh, and Nam Wook Kim. 2021. Toward a unified framework for visualization design guidelines. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–7.
[13] Mengdi Chu, Zefeng Qiu, Meng Ling, Shuning Jiang, Robert S Laramee, Michael Sedlmair, and Jian Chen. 2025. What Makes a Visualization Complex? *arXiv preprint arXiv:2510.08332* (2025).
[14] William S Cleveland and Robert McGill. 1984. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American statistical association* 79, 387 (1984), 531–554.
[15] Michael Correll. 2019. Ethical dimensions of visualization research. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.
[16] Nicholas Diakopoulos. 2018. Ethics in data-driven visual storytelling. In *Data-Driven Storytelling*. AK Peters/CRC Press, 233–248.
[17] Alexandra Diehl, Matthias Kraus, Alfie Abdul-Rahman, Mennatallah El-Assady, Benjamin Bach, Robert Steven Laramee, Daniel Keim, and Min Chen. 2020. Studying visualization guidelines according to grounded theory. *arXiv preprint arXiv:2010.09040* (2020).
[18] Kevin L Elder and Amy A Cesal. 2020. Should We Teach Data Visualization Using Data Visualization Style Guides? *Issues in Information Systems* 21, 4 (2020), 264.
[19] Steven L Franconeri, Lace M Padilla, Priti Shah, Jeffrey M Zacks, and Jessica Hullman. 2021. The science of visual data communication: What works. *Psychological Science in the public interest* 22, 3 (2021), 110–161.
[20] Batya Friedman. 1996. Value-sensitive design. *interactions* 3, 6 (1996), 16–23.
[21] Batya Friedman, David G Hendry, Alan Borning, et al. 2017. A survey of value sensitive design methods. *Foundations and Trends® in Human–Computer Interaction* 11, 2 (2017), 63–125.
[22] Batya Friedman and Peter H Kahn Jr. 2007. Human values, ethics, and design. In *The human-computer interaction handbook*. CRC press, 1267–1292.
[23] Stephen Gale. 1996. A collaborative approach to developing style guides. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 362–367.
[24] Lily W Ge, Yuan Cui, and Matthew Kay. 2023. Calvi: Critical thinking assessment for literacy in visualizations. In *Proceedings of the 2023 CHI conference on human*

factors in computing systems. 1–18.

[25] Tingying He, Petra Isenberg, Raimund Dachselt, and Tobias Isenberg. 2022. BeauVis: A validated scale for measuring the aesthetic pleasure of visual representations. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2022), 363–373.

[26] Jeffrey Heer and Michael Bostock. 2010. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 203–212.

[27] Matthew Honnibal, Ines Montani, Sofie Van Landeghem, and Adriane Boyd. 2020. spaCy: Industrial-strength Natural Language Processing in Python. (2020). doi:10.5281/zenodo.1212303

[28] Jessica Hullman. 2019. Why authors don't visualize uncertainty. *IEEE transactions on visualization and computer graphics* 26, 1 (2019), 130–139.

[29] Jessica Hullman and Nick Diakopoulos. 2011. Visualization rhetoric: Framing effects in narrative visualization. *IEEE transactions on visualization and computer graphics* 17, 12 (2011), 2231–2240.

[30] Eser Kandogan and Hanseung Lee. 2016. A grounded theory study on the language of data visualization principles and guidelines. *Electronic Imaging* 28 (2016), 1–9.

[31] Crystal Lee, Tanya Yang, Gabrielle D Inchoco, Graham M Jones, and Arvind Satyanarayan. 2021. 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*. 1–18.

[32] Kylie Lin, Sean Sheng-tse Ru, David N Rapp, Hui Guan, and Cindy Xiong Bearfield. 2025. What Makes a Visualization Visually Complex?. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–7.

[33] Maxim Lisnic, Cole Polychronis, Alexander Lex, and Marina Kogan. 2023. Misleading beyond visual tricks: How people actually lie with charts. In *Proceedings of the 2023 CHI conference on human factors in computing systems*. 1–21.

[34] Leo Yu-Ho Lo, Ayush Gupta, Kento Shigyo, Aoyu Wu, Enrico Bertini, and Huamin Qu. 2022. Misinformed by visualization: What do we learn from misinformative visualizations?. In *Computer Graphics Forum*, Vol. 41. Wiley Online Library, 515–525.

[35] Sheng Long and Matthew Kay. 2024. To cut or not to cut? a systematic exploration of y-axis truncation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–12.

[36] Oen G McKinley, Saugat Pandey, and Alvitta Ottley. 2025. 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*. 1–17.

[37] Andrew McNutt, Gordon Kindlmann, and Michael Correll. 2020. Surfacing visualization mirages. In *Proceedings of the 2020 CHI Conference on human factors in computing systems*. 1–16.

[38] Dominik Moritz, Chenglong Wang, Greg L Nelson, Halden Lin, Adam M Smith, Bill Howe, and Jeffrey Heer. 2018. Formalizing visualization design knowledge as constraints: Actionable and extensible models in draco. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 438–448.

[39] Kenneth R Ohnemus. 1997. Web style guides: who, what, where. In *Proceedings of the 15th annual international conference on Computer documentation*. 189–197.

[40] OpenAI. 2025. GPT-5 Technical Report. https://openai.com/. Accessed: Nov 25, 2025.

[41] Anshul Vikram Pandey, Katharina Rall, Margaret L Satterthwaite, Oded Nov, and Enrico Bertini. 2015. How deceptive are deceptive visualizations? An empirical analysis of common distortion techniques. In *Proceedings of the 33rd annual acm conference on human factors in computing systems*. 1469–1478.

[42] Saugat Pandey, Oen G McKinley, R Jordan Crouser, and Alvitta Ottley. 2023. Do you trust what you see? toward a multidimensional measure of trust in visualization. In *2023 IEEE Visualization and Visual Analytics (VIS)*. IEEE, 26–30.

[43] Paul Parsons. 2021. Understanding data visualization design practice. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2021), 665–675.

[44] Paul Parsons, Colin M Gray, Ali Baigelenov, and Ian Carr. 2020. Design judgment in data visualization practice. In *2020 IEEE Visualization Conference (VIS)*. IEEE, 176–180.

[45] Paul Parsons and Prakash Shukla. 2020. Data visualization practitioners' perspectives on chartjunk. In *2020 IEEE Visualization Conference (VIS)*. IEEE, 211–215.

[46] Md Dilshadur Rahman, Bhavana Doppalapudi, Ghulam Jilani Quadri, and Paul Rosen. 2025. A Survey on Annotations in Information Visualization: Empirical Studies, Applications and Challenges. *IEEE Transactions on Visualization and Computer Graphics* (2025).

[47] Arvind Satyanarayan, Dominik Moritz, Kanit Wongsuphasawat, and Jeffrey Heer. 2016. Vega-lite: A grammar of interactive graphics. *IEEE transactions on visualization and computer graphics* 23, 1 (2016), 341–350.

[48] Sarah Schottler, Uta Hinrichs, and Benjamin Bach. 2024. Practices and Strategies in Responsive Thematic Map Design: A Report from Design Workshops with Experts. *IEEE Transactions on Visualization and Computer Graphics* (2024).

[49] Jonathan Schwabish, Amy Cesal, Alan Wilson, and Maxene Graze. [n. d.]. Data Visualization Style Guidelines. https://www.datavizstyleguide.com/

[50] Vidya Setlur, Michael Correll, Arvind Satyanarayan, and Melanie Tory. 2023. Heuristics for supporting cooperative dashboard design. *IEEE transactions on visualization and Computer Graphics* 30, 1 (2023), 370–380.

[51] Ben Shneiderman. 2003. The eyes have it: A task by data type taxonomy for information visualizations. In *The craft of information visualization*. Elsevier, 364–371.

[52] Drew Skau and Robert Kosara. 2016. Arcs, angles, or areas: Individual data encodings in pie and donut charts. In *Computer Graphics Forum*, Vol. 35. Wiley Online Library, 121–130.

[53] Chase Stokes, Vidya Setlur, Bridget Cogley, Arvind Satyanarayan, and Marti A Hearst. 2022. Striking a balance: Reader takeaways and preferences when integrating text and charts. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2022), 1233–1243.

[54] Chris Stolte, Diane Tang, and Pat Hanrahan. 2002. Polaris: A system for query, analysis, and visualization of multidimensional relational databases. *IEEE Transactions on visualization and computer graphics* 8, 1 (2002), 52–65.

[55] Edward R Tufte and Peter R Graves-Morris. 1983. *The visual display of quantitative information*. Vol. 2. Graphics press Cheshire, CT.

[56] Huichen Will Wang, Kylie Lin, Andrew Cohen, Ryan Kennedy, Zach Zwald, Carolina Nobre, and Cindy Xiong Bearfield. 2025. Do You" Trust" This Visualization? An Inventory to Measure Trust in Visualizations. *arXiv preprint arXiv:2503.17670* (2025).

[57] Zezhong Wang, Shan Hao, and Sheelagh Carpendale. 2024. Card-Based Approach to Engage Exploring Ethics in AI for Data Visualization. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–7.

[58] Colin Ware. 2019. *Information visualization: perception for design*. Morgan Kaufmann.

[59] Kanit Wongsuphasawat, Dominik Moritz, Anushka Anand, Jock Mackinlay, Bill Howe, and Jeffrey Heer. 2016. Towards a general-purpose query language for visualization recommendation. In *Proceedings of the workshop on human-in-the-loop data analytics*. 1–6.

[60] Shishi Xiao, Suizi Huang, Yue Lin, Yilin Ye, and Wei Zeng. 2023. Let the chart spark: Embedding semantic context into chart with text-to-image generative model. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2023), 284–294.

[61] Fernando Yanez, Cristina Conati, Alvitta Ottley, and Carolina Nobre. 2025. The State of the Art in User-Adaptive Visualizations. In *Computer Graphics Forum*, Vol. 44. Wiley Online Library, e15271.

[62] Zehua Zeng and Leilani Battle. 2023. A review and collation of graphical perception knowledge for visualization recommendation. In *Proceedings of the 2023 CHI conference on human factors in computing systems*. 1–16.

[63] Zehua Zeng, Junran Yang, Dominik Moritz, Jeffrey Heer, and Leilani Battle. 2023. Too many cooks: Exploring how graphical perception studies influence visualization recommendations in draco. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2023), 1063–1073.

A The JSON Schema described in Section 3

The categorizations are not mutually exclusive; a single guideline may be placed in multiple sections by the coder. Eg., if a chart-specific guideline cites accessibility as its rationale, it is included under both `chart` and `accessibility`.

```
# ----------------------------------------
# Reusable guideline object (list-of-lists)
# ----------------------------------------
GUIDELINES = {
    "do": [],
    "dont": [],
    "accessibility": [],
    "ethical_considerations": []
}

template = {
    # ========================================
    # Organization-level metadata
    # ========================================
    "organization": {
        "id": "",                    # str: short org code
        "name": "",                  # str: organization name
        "industry": "",              # str: industry label
        "style_guide_version": "",   # str: version if specified
        "last_updated": "",          # str/date: last updated if known
        "url": "",                   # str: canonical URL for the guide

        "audience": [],              # list[str]: intended users (e.g., designers)
        "preferred_tool": []         # list[str]: tools mentioned (e.g., Tableau, BI)
    },

    # ========================================
    # High-level conceptual guidance
    # ========================================
    "fundamentals": {
        "definition": "",            # str: what "good visualization" means
        "principles": [              # list[{"goal","description"}]
            {"goal": "", "description": ""}
        ],
        "glossary": [                # list[{"term","definition"}]
            {"term": "", "definition": ""}
        ]
    },

    # ========================================
    # Chart-type-specific recommendations
    # ========================================
    "charts": [{
        "chart_type": "",            # str: e.g., bar, line, pie
        "avoid": False,              # bool: org discourages / bans this chart type
        "description": "",           # str: when/how to use
        "task_type": [],             # list[str]: tasks explicitly associated with chart type
        "data_type": [],             # list[str]: data structures explicitly associated
        GUIDELINES                   # do/dont/accessibility/ethical_considerations
    }],
```

```
    # ==========================================
    # Cross-chart "chart element" rules
    # (applies regardless of chart type)
    # ==========================================
    "chart_elements": {
        "font": GUIDELINES,
        "title": GUIDELINES,
        "axes": GUIDELINES,
        "gridlines": GUIDELINES,
        "legend": GUIDELINES,
        "interactive_legend": GUIDELINES,
        "layout": GUIDELINES,
        "annotations_narrative": GUIDELINES,
        "iconography": GUIDELINES,
        "tooltips": GUIDELINES,
        "popovers": GUIDELINES,
        "animation": GUIDELINES
    },

    # ==========================================
    # Task definitions and task->chart mappings
    # ==========================================
    "tasks": [{
        "name": "",                    # str: e.g., "comparison", "ranking"
        "definition": "",              # str: what the task means (if present)
        "charts": []                   # list[str]: recommended chart types (optional)
        # NOTE: if charts is non-empty -> chart-level task mapping is present
    }],

    # ==========================================
    # Other guideline sections
    # ==========================================
    "interaction_guidelines": GUIDELINES,   # hover, zoom, filter, etc.
    "general_guidelines": GUIDELINES,       # cross-cutting best practices

    "accessibility_guidelines": {           # overall accessibility section
        "do": [], "dont": [], "ethical_considerations": []
    },

    "color_guidelines": {
        "palettes": [{
            "name": "primary_palette",      # str
            "description": "",              # str
            "colors": [{
                "name": "",                 # str
                "hex": "",                  # str: "#RRGGBB"
                "usage": ""                 # str: "categorical", "background", etc.
            }],
            "guidelines": []                # list[str] or list[guideline objects]
        }],

    }
}
```

B Value Tags

Unified superset of value tags, with definitions and related keywords used for identifying values in organizational visualization style guides.

Glossary of value tags.

Family	Tag	Definition
Accuracy / Integrity	accuracy	faithful data representation without distortion or misleading emphasis.
	honesty	avoids manipulation, exaggeration, or deceptive framing of information.
	reliability	produces consistent, stable interpretation across conditions.
	trust	fosters user confidence in the visualization's integrity and intentions.
	completeness	includes essential context and avoids omission of critical information.
	non_maleficence	avoids causing harm, confusion, or misinterpretation.
Clarity / Efficiency	clarity	easy interpretation with minimal ambiguity or cognitive burden.
	comprehensibility	supports deeper understanding beyond surface-level clarity.
	simplicity	removes unnecessary complexity to emphasize key information.
	efficiency	enables rapid reading or decision-making with minimal effort.
Accessibility / Inclusivity	accessibility	supports users with diverse abilities through inclusive design.
	inclusivity	accommodates diverse identities, backgrounds, and user contexts.
	cultural_sensitivity	respects cultural norms, symbols, and interpretive conventions.
Branding / Coherence	branding	aligns visual style with organizational identity and brand rules.
	coherence	maintains internal stylistic consistency across visual elements.
	authority	conveys professionalism, expertise, or institutional legitimacy.
Action / Narrative	actionability	supports decision-making with clear, actionable insights.
	storytelling	communicates a narrative or guided interpretive structure.
	persuasion	aims to influence attitudes, behaviors, or conclusions.
	emotional_appeal	leverages affective cues to increase engagement or meaning.
Responsibility / Rights	privacy	protects sensitive or identifiable information about individuals.
	autonomy	supports user freedom to explore or interpret independently.
	informed_consent	respects individuals' rights regarding data use or display.
	responsibility	prioritizes ethical use and potential impacts of visualization.
	beneficence	promotes user well-being and positive outcomes.
	sustainability	considers environmental impact or resource efficiency.
	solidarity	supports community cohesion or collective welfare.

Notes: Value families group related tags into broader ethical and functional dimensions used in our coding.