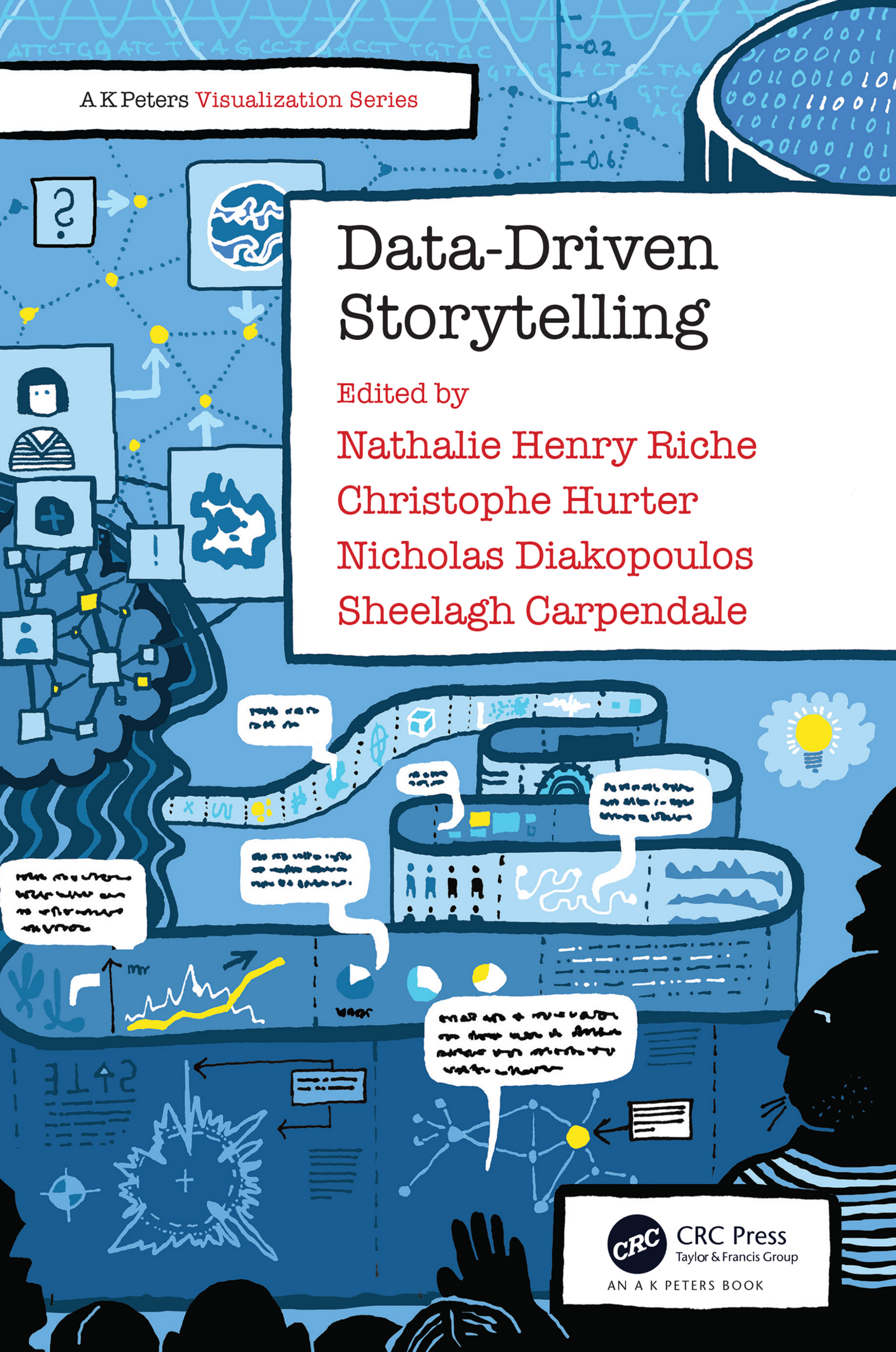
A K Peters Visualization Series

Data-Driven Storytelling

Edited by

Nathalie Henry Riche
Christophe Hurter
Nicholas Diakopoulos
Sheelagh Carpendale



CRC Press
Taylor & Francis Group
AN A K PETERS BOOK

CHAPTER 10

Ethics in Data-Driven Visual Storytelling

Nicholas Diakopoulos

Northwestern University

CONTENTS

The Accountable Journalism Project at the University of Missouri touts having compiled a database of more than 400 codes of media ethics from around the world.* Yet a topical search of the codes tagged with "data

* https://accountablejournalism.org.

journalism" yields only 17 results—that is only about 4%. Well-regarded ethics codes like those from National Public Radio (NPR)* and the Associated Press (AP)† do at least mention data and graphics, and a revised 2017 version of the *AP Stylebook* includes a whole chapter on best practices in data journalism, but, for example, a widely used ethics guide from the Society for Professional Journalists (SPJ) does not mention data at all and includes only an oblique reference not to "distort…visual information."‡ Despite the use of data-driven graphics in the news context for at least a couple hundred years, a robust articulation of ethical mandates for responsible usage has largely failed to emerge and become widely adopted.

This chapter is a step in the direction of articulating a framework for thinking about ethical decisions and pitfalls that relate directly to creating data-driven visual stories. In particular, I approach this within the pragmatist-realist discourse on data visualization (Dick 2016), meaning that I focus on the utilitarian usage of data visualization in the journalistic news media in comparison to other usages for persuasion or for artistic storytelling, which have different ethical commitments.

Ethics is concerned with articulating apt behavior and conduct: how ought one act according to standards of character? Journalism ethics scholar Stephen Ward defines it as "the study and practice of what constitutes the best regulation of human conduct, individually and socially" (Ward 2015). In the context of this chapter, ethics is about making sound editorial decisions along each step of the data-driven visual storytelling process: from the collection and acquisition of data to the analysis and presentation of information.

Principles can be useful guideposts for generally applicable beliefs about ethical behavior. It is important to articulate principles and delineate how they apply because having clearly described codes of ethics can have beneficial impacts on professionals' behavior (Powell and Jempson 2014). Here I do not articulate new principles but draw on those proffered by McBride and Rosenstiel (2013) and examine how they apply to the unique and specific demands of data-driven story production. These principles include:

1. "Seek truth and report it as fully as possible."
2. "Be transparent."
3. "Engage community as an end, rather than as a means."

---

* http://ethics.npr.org/.

† http://www.ap.org/company/News-Values.

‡ http://www.spj.org/ethicscode.asp.

Principle #1 is particularly relevant to data stories because of the many possibilities to mislead, misinform, obfuscate, or even deceive using visualization. As will be discussed in the next sections, opportunities for (mis-) guiding end-users' interpretation of data can arise during every phase of production from acquisition and how the data is quantified, to how it is normalized, aggregated, filtered, visually encoded, annotated, and made interactive.

Principle #2 is relevant to the ways in which insights from data are found and interpreted as stories. Given that many interpretations may arise from a data set, it becomes important for the storyteller to be able to trace and disclose their rationale and process for arriving at a preferred interpretation. This is particularly important as it relates to data acquisition and transformation steps that may go unseen unless they are given explicit consideration for disclosure.

Finally, principle #3 hints at another dimension of importance when telling stories with data: the individual people who the data represents are to be treated with humanity and respect and not be seen only as a means for telling a story. In terms of supporting the goal of community engagement, this principle also connects to interactivity in data-driven storytelling. Designers can engage community through interactivity with the data rather than dictating data as if it is immutable.

Besides moral righteousness, acting ethically in data-driven visual storytelling can have positive outcomes for end-users. Alberto Cairo has addressed ethics in this domain by espousing a utilitarian perspective primarily focused on the end-user benefit of improved understanding and knowledge through the reception of accurate and compelling information (Cairo 2014a, 2017). Acting ethically in this domain entails not only honesty and virtuous intent but also considering and minimizing the potential errors of interpretation that may ultimately mislead viewers (Cairo 2014b; Bradshaw 2015). The utilitarian perspective is worth embracing, and expanding to consider the social benefit of increased trust of individuals or organizations acting ethically by adhering to the principles listed above.

In the next sections I examine ethical considerations that may arise across the data-driven story production pipeline: in data acquisition, in data transformation and interpretation, and in the design and conveyance of those insights for consumption. An understanding of ethical visualization design is further informed by considering how designers can use rhetorical techniques to strategically prioritize certain interpretations

(Hullman and Diakopoulos 2011). The goal is to describe why each ethical consideration is important and relevant to the principles articulated above, providing illustrations from concrete cases where possible. I strive to provide a range of ethical considerations that designers may encounter while acknowledging that this single chapter cannot possibly cover all of the myriad ways in which ethics touches data-driven visual storytelling.

DATA ACQUISITION

Provenance

The origin of the data used, including the possible political or advocacy motives of the data provider are essential to understand when striving to tell informative and truthful data narratives. Who produced the data and what was their intent? Is it complete, timely, and accurate? It is important to provide a baseline of transparency by citing or linking to the data sources used. But there is also the ethical question of whether to use a data set at all, such as if the progenitors of the data may have a stake in misinforming, distracting, or promoting a version of the truth that may be at odds with the public interest, or if the data was obtained illegally such as through hacking or an anonymous leak.

"The wave of bullshit data is rising, and now it's our turn to figure out how not to get swept away," writes Jacob Harris, a former *New York Times* software developer.[*] So-called "PR data" is released by private organizations with the intent of positioning their own organization favorably, or to gain free brand awareness by being listed as a data source on a story that gets a lot of attention. For instance, a story published by *The Washington Post* (and similar stories from many other outlets) used data from the pest company Orkin to present a ranking of the "top 20 rattiest cities" according to where Orkin made the most number of rodent treatments.[†] The numbers suggest the rat problem in Washington, D.C. is bad, worse than New York City even, but it is unclear whether given the bias in the data collection that this is unnecessarily stoking concern for a "rat problem" in the city when it may not really exist. For instance, 2013 data from the U.S. Census American Housing Survey ranked Washington, D.C. as 13th in terms of proportion of occupied housing with reported rat sightings.[‡]

[*] http://www.niemanlab.org/2014/12/a-wave-of-p-r-data/.

[†] https://www.washingtonpost.com/blogs/local/wp/2014/10/13/rats-d-c-calls-pest-company-about-rodents-more-often-than-new-york/.

[‡] http://www.bloomberg.com/news/articles/2015-07-30/these-are-the-most-vermin-filled-cities-in-the-u-s-.

Even for ostensibly reputable sources like governmental or government-funded organizations, it can be important to understand why the data was collected to begin with, how politics may have affected its collection, and whether then it is still suitable to repurpose the data for a given storytelling context. For instance, data journalist Nicolas Kayser-Bril suggests politics may color how the Global Terrorism Database defines and tabulates acts of violence in Eastern Ukraine as terrorism rather than as acts of war between Ukrainian and Russian militaries.*

Quantification

The way in which the world is measured and turned into data is another ethically interesting consideration. This includes not only how the world is sampled and how individual data are measured, but also how measurements are defined to begin with, and where to draw the line between something potentially ambiguous that is counted as category "A" rather than category "B." This can be further complicated when different stakeholders have different definitions. As we put a ruler to the world we cannot help but lose information and context in the creation of data, and so it is important to consider whether what is measured in the data allows for a given story to be faithful to reality. Are we really measuring what we mean to measure?

Whether counting ethnicity, mass shooting events, education, or unemployment, the definition of "what counts" in a particular category matters (Stray 2016). An example of the power of definition comes from the Chicago police department, which from 2010 to 2013 managed to reduce the rate of 8 index crimes by a staggering 56%. But an investigative report from *Chicago Magazine* suggests that these numbers were too good to be true and that one of the ways the stats had been "rinsed and washed" was to misclassify and downgrade offenses so that they were not tallied as part of the eight index crimes.† To tell an honest story about the changes in crime, it is essential to know the exact definitions of what is tracked and how those definitions could be subjectively massaged (or simply misunderstood or misapplied) by the data collector, or be legitimately changed over time. In this case, understanding the social and political pressures on the crime quantification process in Chicago led to an extensive story of its own.

---

* http://blog.nkb.fr/data-free.

† http://www.chicagomag.com/Chicago-Magazine/May-2014/Chicago-crime-rates/.

DATA TRANSFORMATION

Normalization

XKCD comic creator Randall Munroe illustrates the core problem with non-population-normalized mapping in Figure 10.1. The basic issue is that, for many types of data, plotting them on a choropleth map will simply show hot spots in areas of high population. The map ends up showing high population density around urban areas rather than the variable of interest. The correction here is to divide by the population in the area (e.g., state) that is being mapped, thus transforming the variable mapped to a population-normalized value.

Of course, normalization should not be done blithely. The ethical decision here concerns how choices in normalization affect the insight or story that a reader will come away with, including the clarity with which they see that story. For instance, in charting variables that typically have a large exponential skew, such as social media retweets or likes on posts, a

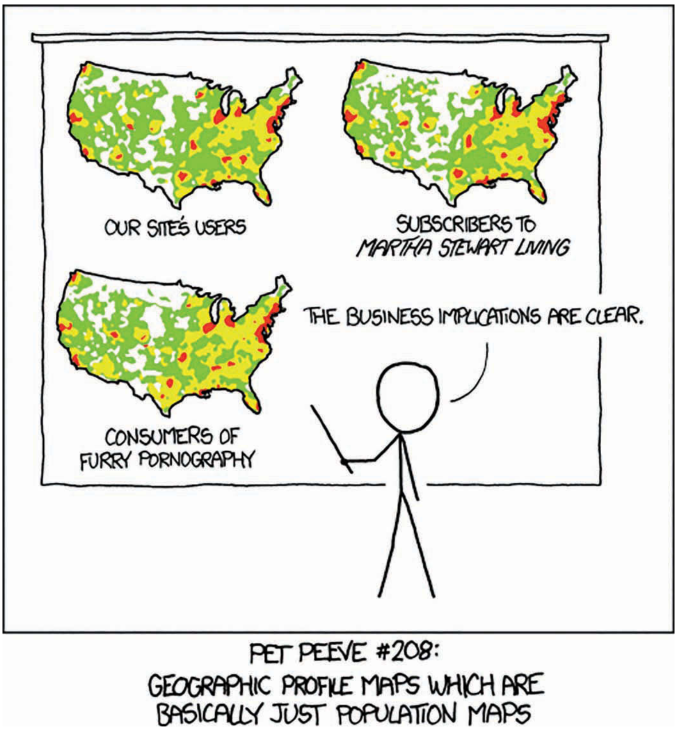


FIGURE 10.1 XKCD comic points out how non-normalized maps often just indicate population density.

visualization designer might choose to log-transform the variable to make the range of values more visible. In charting variables relating to the economy over time it is often necessary to adjust and normalize for inflation so that the underlying signal is faithfully shown.

One way to manage the ethical decision of normalization is to give the end-user more freedom to make this decision on their own terms. Interactivity can facilitate exploration of different types of normalization and their effect on the take-away of a story. An interesting example of this is Google's "Alternative Olympics medal table," which shows the rankings by country of the various Olympic medals in the 2016 games.* At the top of the interface are options to normalize the results by population, gross domestic product (GDP), and other factors to see how the various countries compare. The presentation draws attention to the fact that the rankings will differ when normalized in different ways, adding to the complexity and depth of the data presentation while putting the ethical decision of normalization in the hands of end-users.

Aggregation

When telling a story with aggregated data—data that is derived through mathematical operations or combinations—it is ethically preferred to ensure that a consistent form of aggregation is applied to any data that are to be visually compared. This includes basic methods of aggregating a set of numbers such as taking the mean or median, as well as other operations like binning values to create a histogram. For instance, bin widths for aggregating values into counts for a histogram should ideally be made consistent so that visual comparisons are made according to commensurate aggregations. When aggregating different data measures through operations such as summation, it is important to ensure that the definitions are compatible so that the summation is meaningful and the combined value is easily interpretable.

Consider the chart from *The Economist* in Figure 10.2, which shows historic and forecast vehicle sales from different world geographies. The sharp upward trajectory of the line for China tells a clear story: sales in China are moving up quickly while they are declining or stagnant in North America and Europe. But an important caveat here is that the data for China (and India) has been aggregated differently than for the other geographies depicted—it includes truck, SUV, and minivan sales whereas

* https://landing.google.com/altmedaltable/results/.

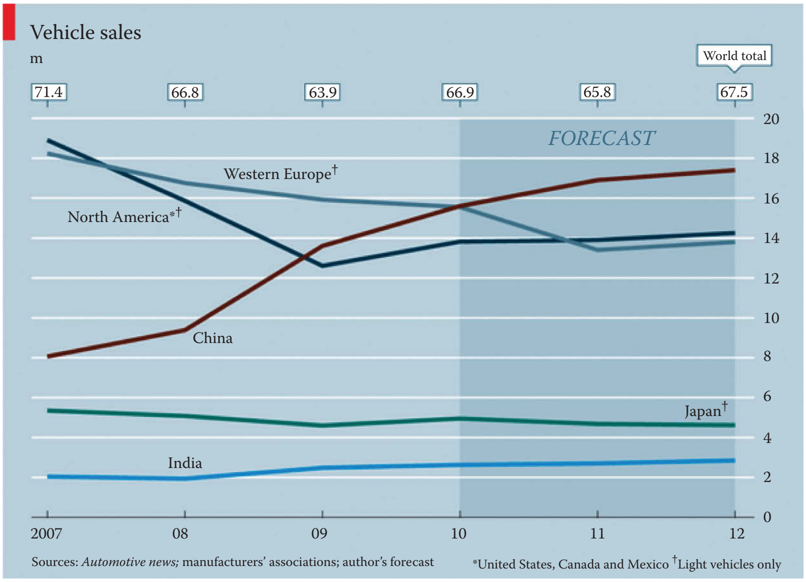


FIGURE 10.2 Historic and forecast vehicle sales around the world. The visual comparison here is misleading because the different data series are defined and aggregated in different ways.

the other geographies exclude those sales and only aggregate light vehicle sales. While the graph does have a small note indicating "light vehicles only" and is transparent about how it has aggregated the data, many readers may simply glance at the chart and see the visual impression of a quickly growing China in comparison to other geographies.

Algorithmic Derivation

In addition to normalization and aggregation, more sophisticated algorithms are sometimes employed to derive new pieces of data. For instance, an analyst might use a classifier to predict a social media user's age or sex based on their connections or posts, or a sentiment analysis routine to determine whether an online review is positive or negative in tone, or a geocoder to translate a written address into a latitude/longitude pair. But any of these algorithmic processes can introduce errors into data and it is important to consider how that might impact the intended story. In 2009, the Los Angeles Police Department was notified of geocoding errors in its public crime mapping portal.* When the geocoder could not locate the

* http://www.latimes.com/local/la-me-geocoding-errors5-2009apr05-story.html.

address of a crime, it would default to an arbitrary location in Los Angeles which made it look like there was a crime hotspot downtown when in fact this was just the result of errors accumulating. When employing algorithmically derived data in visual stories, it is necessary to understand the nature of the errors introduced and try to mitigate the effect those errors might have on end-users' understanding. This can be done by communicating errors or uncertainties in the data via transparency practices such as methods sidebars or by open-sourcing the methods and data (Stark and Diakopoulos 2016).

Filtering

Filtering what is shown is an essential operation that enables successful data-driven storytelling by focusing attention on relevant cases and removing distraction. Filtering data, or the visual process of cropping, requires ethical consideration with respect to the relevance (or irrelevance) of variables, individual cases, or ranges of variables being visualized. In some cases, outliers may be removed from the display (or removed before an aggregation step). Filtering also relates to the idea of cherry-picking—selecting only the data to display that supports an opinion the author has or which otherwise benefits the author in some way. Intentional cherry-picking is surely unethical, but even unintentionally filtering away data that forms essential context or comparison for nonfiltered data could end up being misleading. An ethical question worth engaging is whether the main takeaway from a visual is significantly altered as a result of any filtering that has been applied to the data or to the visual field.

Anonymization

Personally identifying information such as names and addresses is sometimes present in data sets that may underpin a data-driven story. When names and addresses are coupled with other personal information such as political views, arrest or criminal records, or even just consumer behavior (e.g., buying a gun), this can lead to situations where personal privacy issues may come into tension with publishing in the public interest (Bradshaw 2015). While it might be considered an issue of public interest if an official figure like a mayor has an arrest record, it is unclear that having this information for an individual not in the public eye is newsworthy.

Anonymization can be a useful data transformation that protects individual privacy by filtering a data set of any personally identifiable traits that might allow an end-user to infer any identities from the data presentation.

In some cases, anonymization may trigger a need for a change in granularity of aggregation or visual presentation. For instance, consider a map published in 2012 by *The Journal News* in Westchester, NY showing dots at each address where there was a gun permit holder. After publication there was considerable pushback and concern that the map made it too easy to see personal information that could create risks for individuals. Because members of a community can easily infer the identity of each dot based on the address shown, complete anonymization would require a reduction in visual fidelity such as by visualizing the data aggregated by zip code.*

CONVEYING AND CONNECTING INSIGHTS

Visual Mapping and Representation

Edward Tufte's advice on graphic integrity still rings true: "A graphic does not distort if the visual representation of the data is consistent with the numerical representation" (Tufte 2001). There are a number of ways in which data can be mapped to distort and mislead the viewer, and which the ethical communicator must learn in order to avoid.

The way in which data is mapped spatially along a set of axes is an important decision in conveying an insight. Set the range of an axis too small and a trend or spike could get squashed, effectively hiding it from the viewer. Set the range too high and that same trend or spike could give an exaggerated and disingenuous impression. Sometimes it is appropriate to start an axis at zero (e.g., for bar charts because values are read according to the length of the bar), whereas in other cases (e.g., for dot plots or line charts), a nonzero axis starting point is acceptable and can even help clarify nuances to the data. In addition to truncating an axis, inverting an axis is another mechanism that can mislead (Pandey et al. 2015). By flipping an axis, a design might cause an end-user to read an increase as a decrease and vice versa.

Figure 10.3 shows an example of a truncated y-axis that is misleading. In moving from a value of 75% in 2008–2009 to 78% in 2009–2010, the height of the bar (represented as a stack of books) is doubled, visually exaggerating the actual increase. Another misleading aspect of this chart is that the height of the bar from 78% to 79% goes up quite a bit, but for the exact same increase from 81% to 82%, the bars are barely different in

* http://www.poynter.org/2012/where-the-journal-news-went-wrong-in-publishing-names-addresses-of-gun-owners/199218/.

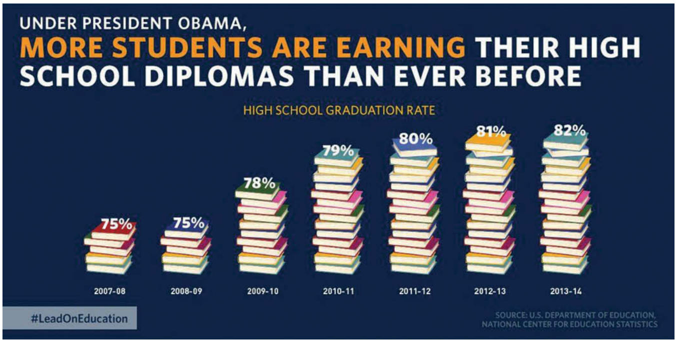


FIGURE 10.3 A misleading chart tweeted by the White House uses a truncated y-axis and inconsistent mapping between bars.

height. In both cases, the height of the bar should go up the same amount so that the data to visual mapping is always consistent.

Other choices in visual mapping and representation can also mislead viewers. A user study of deception techniques in data visualization showed that stretching a chart and changing its aspect ratio can influence how severely a trend is perceived (Pandey et al. 2015). A classic mechanism for deception is to map a value to a 2- or 3-dimensional shape, which distorts the representation because as the value increases linearly the area or volume of a 2- or 3-dimensional shape will grow more quickly and thus be more visually prominent than perhaps warranted (Huff 1954).

Implied Relationships

Plotting two variables in the same visual field, such as with shared axes, can lead end-users to see relationships, associations, or correlations between those variables. Tyler Vigen runs a blog called “Spurious Correlations” (now a book) which draws attention to the absurdity of a range of unlikely statistical correlations. For instance, a chart taken from that blog shown in Figure 10.4 plots the number of people who drowned by falling into a pool versus the number of films that Nicolas Cage has appeared in. It visually indicates that the two lines have a good deal of correlation (there is also statistical correlation with Pearson $r = 0.67$). There’s no explanation and no real reason for why these two variables should be related, yet such associations can easily be implied visually even when there is no logical rationale for a connection.

While such charts like Figure 10.4 are crafted to make light of how we often find meaningless relationships in big data, it is not uncommon to find published charts that imply an association between variables which

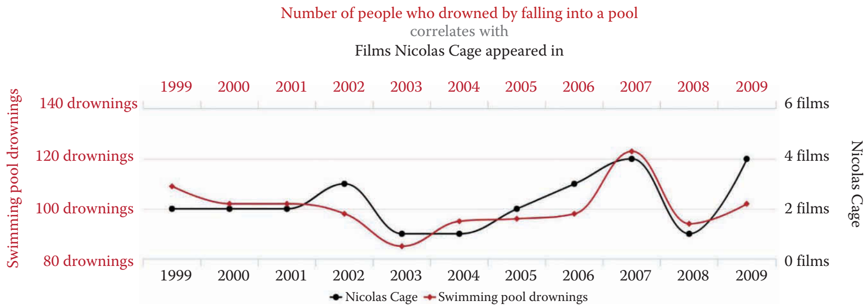


FIGURE 10.4 An absurd chart showing a spurious correlation that misleadingly connects two unrelated variables.

the human mind could easily interpret as causal. Take for instance the chart shown in Figure 10.5, which indicates the annual GDP growth in the United States according to whether there was a Democratic or Republican president in office. Simple looking at the top two bars in that graph visually shows that the economy *has* done better when there was a Democrat as president. But, thankfully, *The Economist* spends the next 2 minutes of the videographic from which this chart was excerpted explaining that in fact this association is not valid or causal and that research shows there are no factors that explain it other than lucky timing.

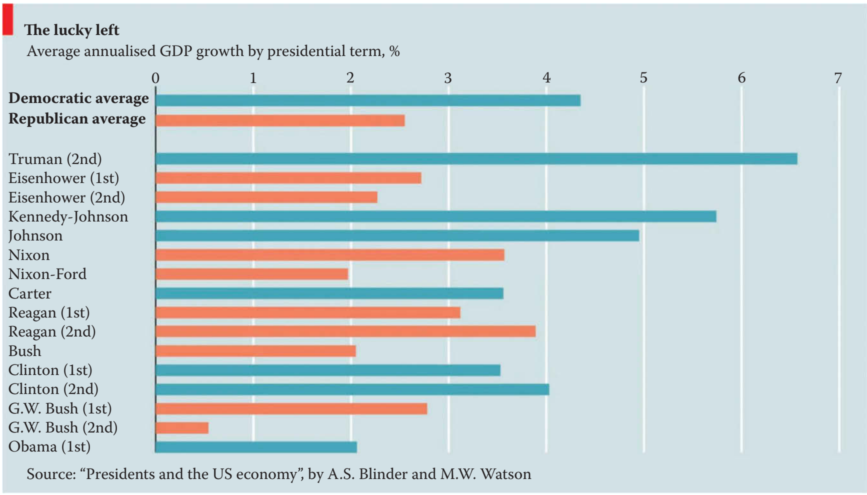


FIGURE 10.5 This chart suggests an association between the president in office and the robustness of the economy, but the rest of the videographic debunks this relationship by explaining the context around each president's term.

Another way in which relationships can be implied is through the sequencing and ordering of views within a narrative visualization (Hullman and Diakopoulos 2011; Hullman et al. 2013b). Temporal, spatial, or general to specific comparisons are often conveyed through subsequent views in interactive slideshow narratives or videographics; causal interpretations of such comparisons can result (Hullman et al. 2013b). Definitions of narrative revolve around the notion of a sequence of events that are causally related (Segel and Heer 2010). The human mind is prone to filling in causal explanations when we experience a story (Gottschall 2013), so it is particularly important to be aware of how readers may form interpretations of variable associations and if possible, to preempt common misinterpretations.

The overarching ethical point here is that data-driven storytellers must be careful that the variable relationships communicated in visualizations are not bogus or suggestive in a way that may be misleading. Whether those variable relationships are implied in the same visual space and chart, or whether they are implied through the sequencing and temporal juxtaposition of views, it is important to only convey those variable relationships (and their causal nature) when there is sufficient and reliable accompanying evidence, such as may come from additional research, context, or statistical analysis.

Context and Annotation

The word "data" derives from a Latin term meaning "something given." It is therefore helpful to remember that data are not facts per se, but are "givens" from which a variety of interpretations can be derived. The meaning end-users impute from data is heavily moderated by how it is presented in relation to other data and to relevant context that sets the stage for interpretation. In visual storytelling, this is often accomplished with thorough labeling and textual explanation or annotation. Text provides that vital bit of context layered over the data that helps the audience come to a valid interpretation of what it really means. Annotations can emphasize, highlight, or prioritize particular pieces of data or interpretations, aiding the storytelling process by directing attention or preempting the user's curiosity on seeing a salient outlier, aberration, or trend (Hullman et al. 2013a).

One area that can be particularly fraught is in choosing labels for categories of people, as this inherently takes a stance on the existence and relevance of that category to the story, despite possible controversy over definitions or boundaries. This came up in a *New York Times* graphic in

2015 when a table representing lawmakers' support about the Iran nuclear deal included a column labeled "Jewish?"* The graphic was subsequently changed by removing the column, but not before provoking a reaction in which several viewers pointed out the irrelevance of the label "Jewish" in this political context.

Ethically, the goal in providing context and annotation is to provide context that is relevant, does not distract, and guides the reader towards the most plausible, logical, and faithful interpretation of the data. If a particular editorial interpretation of the data is conveyed via the context and annotations, then ideally this would also be indicated to the end-user. Care should be taken so that labels, legends, definitions, and other context are presented in a straightforward way that minimizes the potential for misinterpretation and which is complete enough for the viewer to fully comprehend the interpretation presented.

Interactivity

Interactive visualizations provide capabilities for end-users to explore data and have agency as they move through the story presented. Designers can develop a dialogue with users rather than dictate a unitary interpretation that could convey a false sense of finality. Users can navigate, search, and filter the data according to their own interests, allowing for a greater degree of relevance and personalization. Some of the most consequential decisions that data-driven story designers make when creating interactive visualizations relate to the *defaults*. Default views, default or suggested search terms, and default parameters such as slider values, all heavily bias end-users' interaction patterns. For some users, all they will see are the defaults. Ethically, it is important to choose defaults that are well-researched, based in evidence, and not arbitrarily chosen to provide an enticing or misleading view. *The New York Times* "Better to Rent or Buy?"† visual calculator does this well: the default inflation rate used by the calculator is 2.0%, which is right in the ballpark of values for the Consumer Price Index over the last few years.‡ Besides defaults, other goal directions can also be embedded via interactivity such as by making suggestions for what the user should search for, or prompting the user to explore in particular ways rather than explore more freely (Hullman and Diakopoulos 2011).

* http://publiceditor.blogs.nytimes.com/2015/09/11/iran-deal-graphic-jewish-lawmakers-was-insensitive/.

† http://www.nytimes.com/interactive/2014/upshot/buy-rent-calculator.html.

‡ http://data.bls.gov/timeseries/CUUR0000SA0L1E?output_view=pct_12mths.

TABLE 10.1 Ethical Considerations across Different Phases of Data-Driven Storytelling

Data Acquisition	Data Transformation	Conveying and Connecting Insights
• Provenance • Quantification	• Normalization • Aggregation • Algorithmic derivation • Filtering • Anonymization	• Visual mapping and representation • Implied relationships • Context and annotation • Interactivity

SUMMARY

This chapter has enumerated a range of ethical factors (see Table 10.1) that are deserving of careful deliberation across the visual-data storytelling process. By considering tradeoffs in choices related to these various factors, the intent of this chapter is to help the data storyteller become more cognizant of their role and responsibility in guiding the interpretation of the audience. Whether in the data acquisition or transformation stages, when mapping data to charts, or connecting those charts into a sequence, ethics in this domain is predominantly about making sound editorial decisions that ensure the reception of accurate interpretations of the data. Of course, whether a story is interpreted accurately is also contingent on the data and visual literacy of the audience to begin with, and so ethical storytelling also attempts to understand the audience (see Chapter 9 entitled “Communicating Data to an Audience”) and provide scaffolding of necessary literacies to ensure accurate interpretation. When errors do arise and there is a widespread unintended reception or interpretation that is untrue or misleading, the ethical storyteller will issue a correction, and update the story so that any misinterpretation can be avoided in the future. Striving to show the truth, being transparent, and engaging the community as an end is an ongoing endeavor in mindfulness, but one worth investing in to build a more-trustworthy craft of honest visual data storytelling.

ACKNOWLEDGMENTS

This chapter benefitted greatly from the initial brainstorm and discussions on the topic with Gordon Bolduan, Marian Dörk, Kennedy Elliott, and Xaquín G.V. at the Dagstuhl seminar.

REFERENCES

Bradshaw, Paul. 2015. “Data Journalism.” In *Ethics for Digital Journalists: Emerging Best Practices*, edited by Lawrie Zion and David Craig. Routledge.

Cairo, Alberto. 2014a. “Ethical Infographics: In Data Visualization, Journalism Meets Engineering.” *IRE Journal*, Spring 2014 Issue: 25–27.

———. 2014b. "Graphic Lies, Misleading Visuals." In *New Challenges for Data Design*, edited by Davic Bihanic, 103–16. Springer-Verlag, London.

———. 2017. "Moral Visualization." In *Information Design: Research and Practice*, edited by Sue Walker, Alison Black, Paul Luna, and Ole Lund. Routledge, Abingdon, UK.

Dick, Murray. 2016. "Developments in Infographics." In *The Routledge Companion to Digital Journalism Studies*, edited by Bob Franklin and Scott II Eldridge. Routledge, London.

Gottschall, Jonathan. 2013. *The Storytelling Animal: How Stories Make Us Human*. Mariner Books, New York.

Huff, Darrell. 1954. *How to Lie with Statistics*. Norton, New York.

Hullman, Jessica, and Nicholas Diakopoulos. 2011. "Visualization Rhetoric: Framing Effects in Narrative Visualization." *IEEE Transactions on Visualization and Computer Graphics* 17 (12): 2231–40.

Hullman, Jessica, Nicholas Diakopoulos, and Eytan Adar. 2013a. "Contextifier: Automatic Generation of Annotated Stock Visualizations." In *Proceedings of the Human Factors in Computing Systems (CHI)*.

Hullman, Jessica, Steven Drucker, Nathalie Henry Riche, Bongshin Lee, Danyel Fisher, and Eytan Adar. 2013b. "A Deeper Understanding of Sequence in Narrative Visualization." *IEEE Transactions on Visualization and Computer Graphics* 19 (12): 2406–15.

McBride, Kelly, and Tom Rosenstiel. 2013. "The New Ethics of Journalism : Principles for the 21st Century." In *The New Ethics of Journalism: Principles for the 21st Century*, edited by Kelly McBride and Tom Rosenstiel. CQ Press, Thousand Oaks, California.

Pandey, Anshul Vikram, 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 Conference on Human Factors in Computing Systems*, 1469–78.

Powell, Wayne, and Mike Jempson. 2014. "More Accountability in the Digital Age? The Influence of New Technologies." In *Journalists and Media Accountability: An International Study of News People in the Digital Age*, edited by Susanne Fengler, Tobias Eberwein, Gianpietro Mazzoleni, Colin Porlezza, and Stephan Russ-Mohl 115–28.

Segel, Edward, and Jeff Heer. 2010. "Narrative Visualization: Telling Stories with Data." *IEEE Transactions on Visualization and Computer Graphics* 16 (6): 1139–48.

Stark, Jennifer, and Nicholas Diakopoulos. 2016. "Towards Editorial Transparency in Computational Journalism." In *Proceedings of Computation + Journalism Symposium,* Palo Alto, CA, September 30–October 1, 2016.

Stray, Jonathan. 2016. "Quantification." In *The Curious Journalist's Guide to Data*. Tow Center for Digital Journalism, New York.

Tufte, Edward. 2001. "Graphical Integrity." In *The Visual Display of Quantitative Information*, 2nd ed. Graphics Press, Cheshire, Connecticut.

Ward, Stephen J. A. 2015. *Radical Media Ethics: A Global Approach*. Wiley-Blackwell, Chichester, UK.