Purva Kapshikar

10 May 2020

Human Contexts and Ethics of Data…Visualizations

As data visualizations are being increasingly incorporated into news coverage and shared on social media platforms, improper creation or usage of these graphics can cause public misconceptions about critical, sensitive or even polarizing events, such as those pertaining to politics or world affairs. For instance, the following are two graphs showing the same data: annual global temperature from 1880 to 2012 in Fahrenheit.

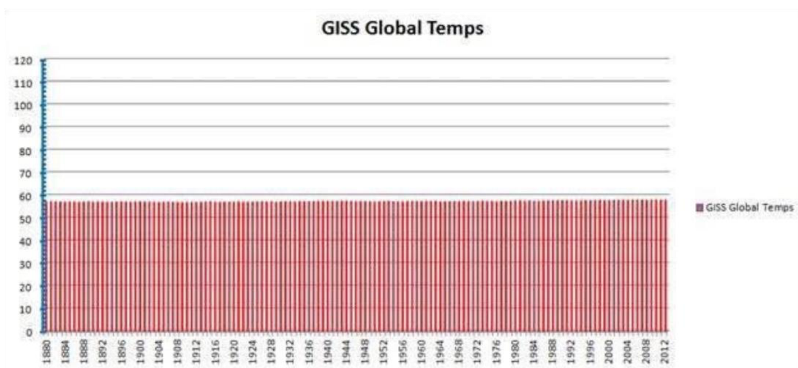

Figure 1: Global temperature data on graph with misleading axis, as shared on Twitter.[1]

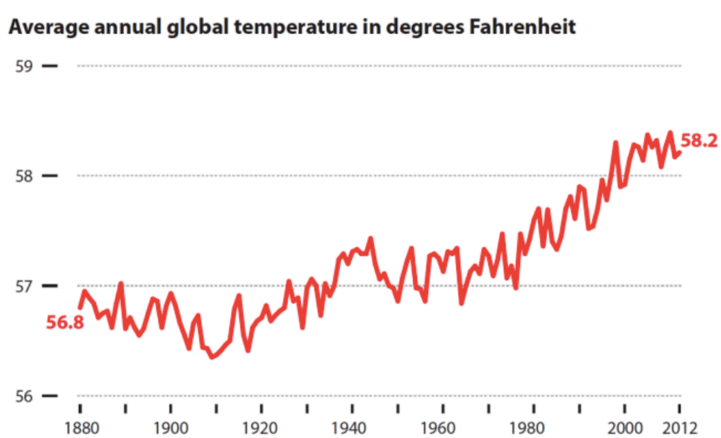

Figure 2: Global temperature data on graph with more appropriate axis.[2]

[1] @EcoSenseNow (Patrick Moore).

[2] "How the public misinterprets data visualizations: and what we can do about it." ESS Data visualisation Milan 2019. Slide 35.

In the former graph, the two-degree difference in temperature over the years is not visible, due the poor choice of scale for the y-axis. It suggests a stability in the climate that does not exist, as even small changes in temperature correspond to "enormous changes in the environment"; thus, the graph is misleading and consequently unethical.[3]

Data visualizations can be effective, as they are more digestible than other forms of data. The point they try to convey is usually very apparent, however, this can be problematic when the data is inaccurate or the visualization is misleading. Data visualizations can improve understanding of a topic if used appropriately. But something that is misleading, whether intentional or not, creates incorrect perceptions or understandings, which can be problematic when acted upon or further disseminated with others.

Utilitarianism is an example of a consequentialist moral theory focused on maximizing well-being through morally good actions.[4] In the context of visualizations, ethical data visualizers should strive to increase their viewers' understanding of topics by including relevant, accurate data. Increased understanding is positively correlated with personal well-being and making better, informed decisions.[5] Thus ethical data visualizers have the responsibility to create good graphics in order to enable comprehension and bring attention to relevant matters.[6]

---

[3] "The Effects of Climate Change." NASA: Global Climate Change, Vital Signs of the Planet.

[4] Boenig-Liptsin, Margo and and Ari Edmundson. "Lecture 11: Aiming at the Good Life With Data." History C184D: Human Contexts and Ethics of Data.

[5] "Chapter 5: Ethics." MSIS 2629 "Data Visualization."

[6] "Ethical Infographics: In data visualization, journalism meets engineering." The Functionary Art.

Similar to data, data visualizations can come across as objective when they have underlying biases and motivations. While it may be tempting to consider visualization as ethically neutral because it involves collecting and reporting facts, "the process of observing the world and quantifying it is a political act, and deserves ethical consideration."[7] Consequently, journalists and designers need to include measures of uncertainty in their work, such as margins of error, when appropriate. The data deluge seems to be accelerating in all aspects: visualizations make this overwhelming volume, variety and velocity of information more comprehensible, and so they play a larger role in how we get news today. Given the prevalence of fake news, we have to be careful with the information we may store and share. People who come across these visualizations may not have the time, knowledge or experience to ascertain whether they are misleading, and with all the channels of communication accessible to us today, the "power to transmit information electronically multiplies the consequences of irresponsible behavior."[8]

Data visualizations can be misleading if the data itself is misleading. For instance, the wrong statistical tests might be used, certain data – such as those reflecting negative results – might be omitted, or data dredging could occur.[9] But visuals can be misleading as well. Visualization deception usually occurs at two levels: either the chart or the message is interpreted incorrectly.[10] The deceptive techniques cause a message exaggeration,

---

[7] Correll, Michael. "Ethical Dimensions of Visualization Research." Page 2.
[8] Forum Guide to Data Ethics. National Forum on Education Statistics, U.S. Department of Education. Page 2.
[9] Marco, Catherine and Gregory Luke Larkin. "Research ethics: ethical issues of data reporting and the quest for authenticity." Academic Emergency Medicine.
[10] O'Brien, Shaun. "Interpretations of Data in Ethical vs. Unethical Data Visualizations." Arizona State University. Page 10.

understatement or reversal.[11] The following are two examples of these misleading tactics and techniques.

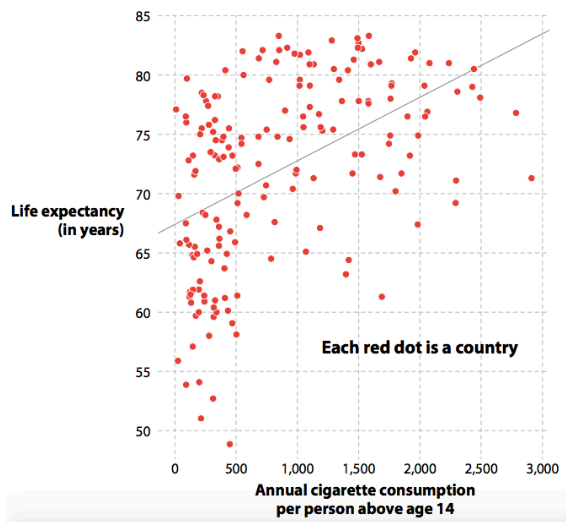

Figure 3: Chart comparing smoking and life expectancy.[12]

The chart above displays a positive correlation between smoking and life expectancy. However, a cause-and-effect relationship cannot be concluded between two variables simply because there is strong correlation between them – namely, correlation does not mean causation.[13] Furthermore, incorrect verbal descriptions can bias understanding of a chart. This chart with the caption "Smoking cigarettes can help you live longer" would take advantage of this correlation to imply a viewpoint. The chart below, however, displays more information that can clarify that: the additional variable of income should be included in the description as well as the chart, otherwise, someone who skims the words without looking closely at the points would be misled. Furthermore, after visually separating the categories, as below, the

---

[11] Ibid.
[12] "How the public misinterprets data visualizations: and what we can do about it." ESS Data visualisation Milan 2019. Slide 54.
[13] "Chapter 5: Ethics." MSIS 2629 "Data Visualization."

amalgamation paradox is evident: the trend that appears in these three separate charts is different from the trend in the chart above with the groups combined.

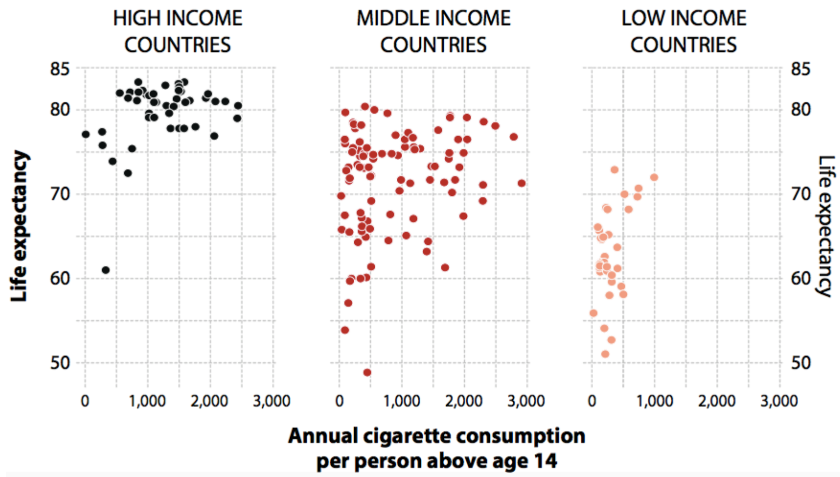

Figure 4: Chart comparing smoking, life expectancy and income.[14]

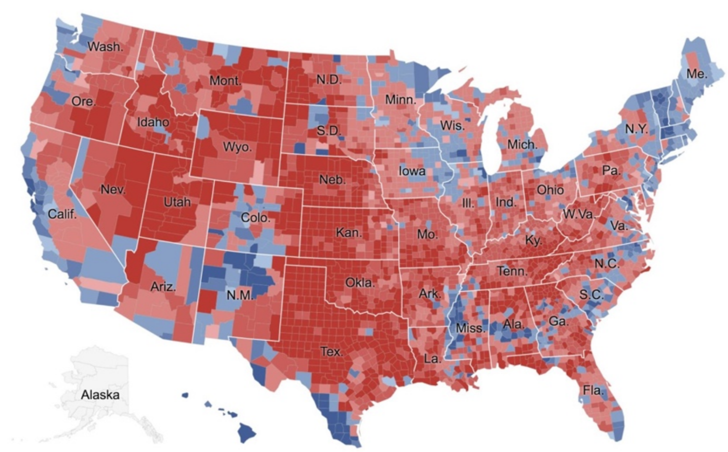

Figure 5: County-level map for 2016 presidential election.[15]

Above is a county-level election map for the 2016 presidential election.[16] Counties that voted Republican do take up a larger surface area of this map – around 80 percent – but this is

[14] "How the public misinterprets data visualizations: and what we can do about it." ESS Data visualisation Milan 2019. Slide 56.
[15] "Election 2016." *The New York Times*.
[16] "How the public misinterprets data visualizations: and what we can do about it." ESS Data visualisation Milan 2019. Slide 25.

due to larger counties that cover more land area. This map was frequently shared on social media, captioned with a variant on the phrase “Try to impeach this”, as if the red area represented a larger number of voters or greater population density.[17] This is an example of taking variables out of context. Instead, the following map, which considers individual voters, is more accurate. It also emphasizes how certain counties – often those that voted Republican – actually have very low population density, clarifying the ambiguity.

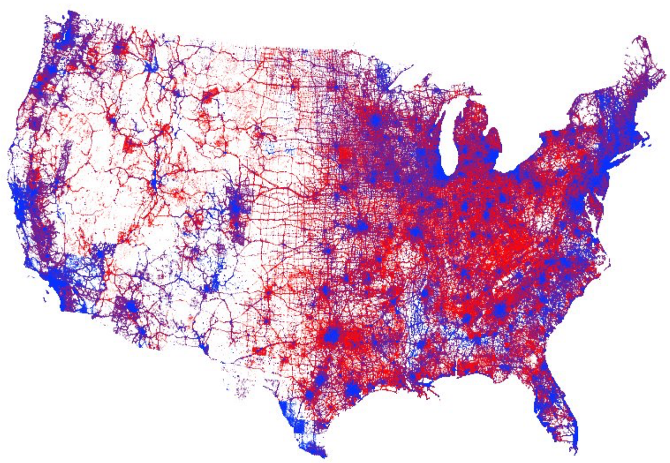

Figure 6: More accurate map for 2016 presidential election, showing location of individual voters.[18]

Including different visuals to display the same data, such as with the data adjusted for a certain parameter, can also increase comprehension. Compared to the county map which can be used out of context, the following graphics explicitly show the same data in different ways, so that a reader can recognize the potential for deception before coming to their own conclusions.

---

[17] @LaraLeaTrump (Lara Trump).
[18] @kennethfield (Kenneth Field).

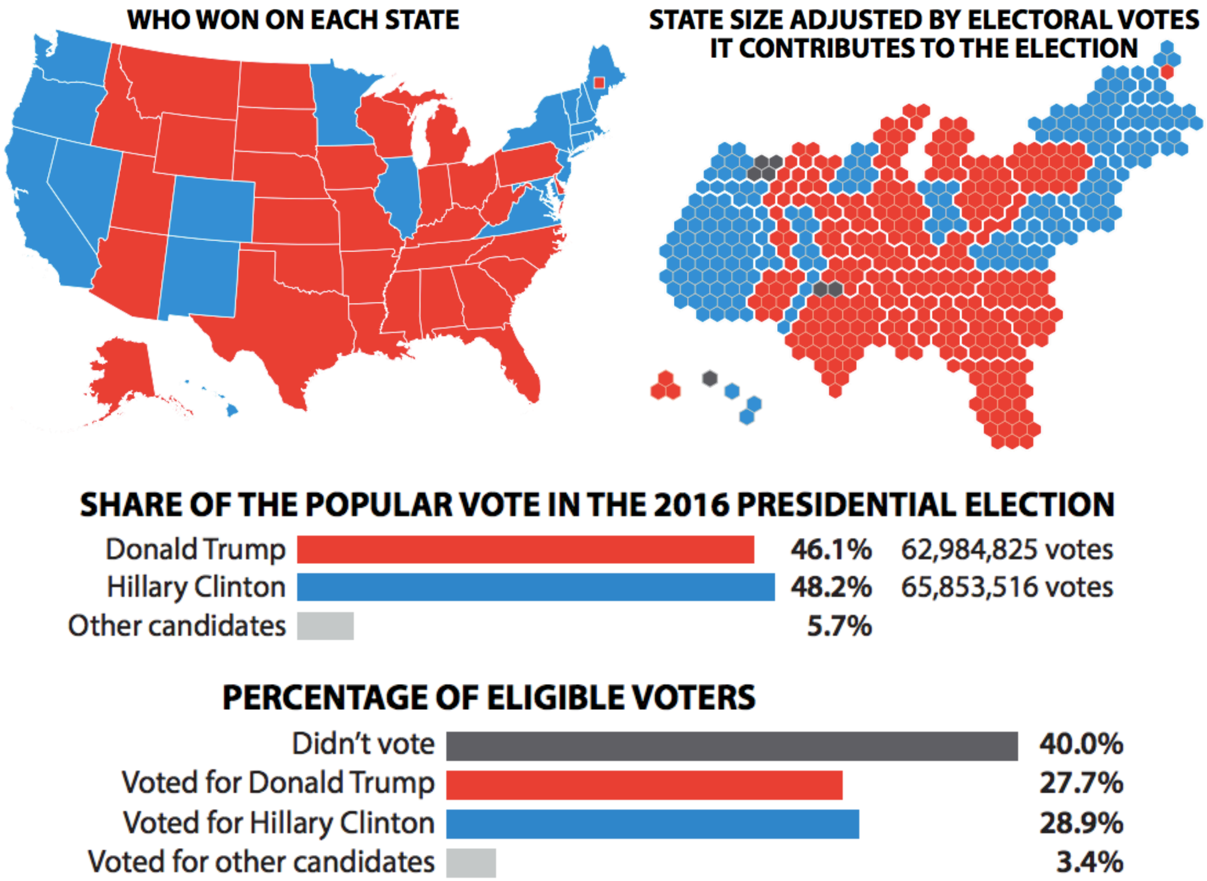

Figure 7: 2016 presidential election maps displaying other variables.[19]

Data visualization does have the potential to be a universal language.[20] To achieve that, visualizers should be careful with not only their data collection and analysis, but also their choice and display of visualization. To increase accessibility and comprehension, they can ask non-experts to review a graphic to see if it could be misconstrued, or even consider adding a component to their graphics that explains how to read or use the visualization, to avoid misunderstandings.[21] Viewers and readers can familiarize themselves with various types of

[19] "How the public misinterprets data visualizations: and what we can do about it." ESS Data visualisation Milan 2019. Slides 28-29.
[20] "How the public misinterprets data visualizations: and what we can do about it." ESS Data visualisation Milan 2019. Slide 6.
[21] "How the public misinterprets data visualizations: and what we can do about it." ESS Data visualisation Milan 2019. Slide 64.

visualizations and ways they might be misleading, such as those explained above. Through experience, they will be better equipped to analyze and understand data visualizations and be more aware of deceptive tactics, which will help them come to more accurate conclusions and avoid sharing misleading visualizations or opinions with others. With both visualizers and readers working to bridge the gap, data visualizations will become more prevalent, more effective and – ideally – consistently ethical.

Bibliography

Acharya, Govind and Hunter Whitney. “Practicing Good Ethics in Data Visualization.” University of California, Davis. 2017. https://www.coursera.org/lecture/dataviz-design/practicing-good-ethics-in-data-visualization-YqHHo.

Boenig-Liptsin, Margo and Ari Edmundson. “Lecture 11: Aiming at the Good Life With Data.” History C184D: Human Contexts and Ethics of Data. 14 February 2020.

Cairo, Alberto. “Ethical Infographics: In data visualization, journalism meets engineering.” IRE Journal. 19 June 2014. https://www.dropbox.com/s/pqgmg02yz0pgju4/EthicalInfographics.pdf.

“Chapter 5: Ethics.” MSIS 2629 “Data Visualization.” 10 June 2019. https://mschermann.github.io/data_viz_reader/ethics.html.

Correll, Michael. “Ethical Dimensions of Visualization Research.” 19 Dec 2018. https://arxiv.org/pdf/1811.07271.pdf.

@EcoSenseNow (Patrick Moore). “Here’s what the dramatic rise in global temperature looks like when a realisitic scale is used on the Y-axis. It is so not out of the ordinary it is completely ridiculous to make up fake stories about CO2 being the cause of the ‘climate crisis’. The climate is very stable.” *Twitter*. 23 April 2019, 11:26 a.m., https://twitter.com/EcoSenseNow/status/1120755895149957130.

“The Effects of Climate Change.” NASA: Global Climate Change, Vital Signs of the Planet. https://climate.nasa.gov/effects/.

“Election 2016.” *The New York Times*. 9 August 2017. https://www.nytimes.com/elections/2016/results/president.

"Ethical Infographics: In data visualization, journalism meets engineering." The Functionary Art. 19 June 2014. http://www.thefunctionalart.com/2014/06/infographics-data-and-visualization.html.

Forum Guide to Data Ethics. National Forum on Education Statistics, U.S. Department of Education. 2010. https://nces.ed.gov/pubs2010/2010801.pdf.

"How the public misinterprets data visualizations: and what we can do about it." ESS Data visualisation Milan 2019. https://ec.europa.eu/eurostat/cros/system/files/alberto_cairo_0.pdf.

@kennethfield (Kenneth Field). "Fair bit of data wrangling but finally...a 1 dot = 1 vote dasymetric dot density 2016 Presidential election map from @ArcGISPro. 65,844,61 blue dots. 62,979,636 red dots. Count 'em! (note: this is just a rough screengrab, will webify & make pretty)" *Twitter.* 5 March 2018, 5:04 p.m., https://twitter.com/kennethfield/status/970827334038237184.

@LaraLeaTrump (Lara Trump). *Twitter*. 28 September 2019. 12:36 p.m. https://twitter.com/LaraLeaTrump/status/1178030815671980032.

O'Brien, Shaun. "Interpretations of Data in Ethical vs. Unethical Data Visualizations." Arizona State University. December 2017. https://repository.asu.edu/attachments/194100/content/OBrien_asu_0010N_17523.pdf.

Marco, Catherine and Gregory Luke Larkin. "Research ethics: ethical issues of data reporting and the quest for authenticity." Academic Emergency Medicine. 28 June 2008. https://doi.org/10.1111/j.1553-2712.2000.tb02049.x.

Pandey, Anshul Vikram, et al. “How Deceptive are Deceptive Visualizations?: An Empirical Analysis of Common Distortion Techniques.” Association for Computing Machinery. 18 April 2015. https://dl.acm.org/doi/10.1145/2702123.2702608.