Reasoning for Data Visualization

How to make good choices

Alberto Cairo

www.albertocairo.com Twitter: @albertocairo
Why do cats & dogs ...?
The Art of Insight:
How Great Visualization Designers Think
(working title)
Wiley, Fall of 2021 or Spring of 2022
A visualization isn’t just an “image”.
A visualization is an argument, or part of an argument.

Therefore, visualizations can’t be designed based just on our personal preferences—although these are important—but on reasoned, justified choices.

**Visualization is similar to writing:** beyond some conventions and constraints regarding symbols, visual grammar, perception, and cognition, visualization can’t be based on “rules” that are set in stone.

Each visualization is unique.
“Facts give us reasons [...] when they count in favor of our having some belief or desire, or acting in some way.”

Derek Parfit, *On What Matters*
Reasoning about visualization. Key questions:

1. Why to visualize?
2. What to visualize?
3. Who to visualize for?
4. How much to visualize?
5. How to visualize it?
6. What style to use?
7. What words to add?

Note: this is just a tentative list of questions
1. Why should my visualization exist?

Do the potential benefits of designing my visualization outweigh the possible harm it might cause?
"Where are the gun permits in your neighborhood?" That’s the question posed by The Journal News, a New York newspaper that published a Google map on Sunday that shows the names and addresses of pistol or revolver permits in Westchester and Rockland counties.

We felt sharing information about gun permits in our area was important in the aftermath of the Newtown shootings.

Janet Hasson, president and publisher of the Journal News Media Group
Why should this data be made public?
Why should it be made public through a map?
Why should it be this type of map?
Even if we decided that this data is worth publishing, wouldn’t a different map be better?

What are the potential consequences of my decisions?
Are the benefits worth the risk of harm?
Ten Considerations Before You Create Another Chart About COVID-19

To sum it up — #vizresponsibly; which may mean not publishing your visualizations in the public domain at all.

As a public health professional, might I ask:

Please consider if what you’ve created serves an actual information need in the public domain. Does it add value to the public and uncover new information?

If not, perhaps this is one viz that should be for your own use only.

We need to go beyond the ethics of journalism and explore the ethics literature in statistics, data science, technology, engineering, etc. because data journalism exists at the intersection of—and borrows from—all those disciplines.
2. What to visualize?

Do I understand my data, its limitations, uncertainty, or glitches?
What or who is being measured (or not being measured) and why?
AT SCHOOL WITHOUT A ROOF

In Florida more than 71,000 students are homeless. During the last decade, this population rocketed as a result of the recession and how hard it has become for the poorest families to find affordable housing.

http://www.lmelgar.me/without-a-roof/
Kansas is the nation’s porn capital, according to Pornhub

Blue states watch more porn. But what’s the matter with Kansas?

According to Pornhub Insights, Kansas leads the nation in porn pageviews per capita at roughly 194. They don’t specify what interval this is over (monthly, weekly, etc), but the state-by-state comparison is nonetheless interesting.

Plotting Obama vote share in 2012 versus porn consumption, it looks like blue states consume more porn per capita than red ones. Aside from Kansas - a clear outlier - and Georgia, the remaining top ten per-capita porn consumers are all blue. Similarly, New Mexico and Maine are the only blue states in the bottom ten per-capita porn consumers.

https://wonkviz.tumblr.com/post/82488570278/kansas-is-the-nations-porn-capital-according-to
Kansas is the nation’s porn capital, according to Pornhub

Blue states watch more porn. But what's the matter with Kansas?

According to Pornhub Insights, Kansas leads the nation in porn pageviews per capita at roughly 194. They don’t specify what interval this is over (monthly, weekly, etc), but the state-by-state comparison is nonetheless interesting.

Plotting Obama vote share in 2012 versus porn consumption, it looks like blue states consume more porn per capita than red ones. Aside from Kansas - a clear outlier - and Georgia, the remaining top ten per-capita porn consumers are all blue. Similarly, New Mexico and Maine are the only blue states in the bottom ten per-capita porn consumers.

https://wonkviz.tumblr.com/post/82488570278/kansas-is-the-nations-porn-capital-according-to
Kansas is the nation’s porn capital, according to Pornhub

Blue states watch more porn. But what’s the matter with Kansas?

According to Pornhub Insights, Kansas leads the nation in porn pageviews per capita at roughly 194. They don’t specify what interval this is over (monthly, weekly, etc), but the state-by-state comparison is nonetheless interesting.

Plotting Obama vote share in 2012 versus porn consumption, it looks like blue states consume more porn per capita than red ones. Aside from Kansas - a clear outlier - and Georgia, the remaining top ten per-capita porn consumers are all blue. Similarly, New Mexico and Maine are the only blue states in the bottom ten per-capita porn consumers.

https://wonkviz.tumblr.com/post/82488570278/kansas-is-the-nations-porn-capital-according-to
...and also data’s limitations and uncertainty:

Uncertainty and graphicity:
How should statisticians, journalists, and designers reveal uncertainty in graphics for public consumption?

Articles and papers about communicating and visualizing uncertainty:

https://www.dropbox.com/sh/jk4ginxyai6ylqu/AABvqdyT1lhlyFN9nKNHyX9Ba?dl=0
3. Who to visualize for?

Have I thought about how my intended audience will access my graphic? Can they read it? Will they understand it?
Pew Research Center
Survey 2014

63% of American Adults Can Correctly Read This Chart
Which of the following statements best describes the data in the graph below?

Average number of decayed teeth per person in different countries

Average sugar consumption (grams per person per day)
The impulse many editors will have when seeing this survey will be: “We shouldn’t use scatterplots!”

Wrong! Use them AND explain them.
The language of visualization is constantly expanding, and we should help readers to expand their own, personal visual vocabularies, as well.

But if you use an unusual chart, EXPLAIN HOW TO READ IT.

https://twitter.com/AlbertoCairo/status/1236773377865658370
We should also strive for accessibility

accessible data viz is better data viz

JUNE 26, 2018 BY COLE NUSSBAUMER KNAFLIC IN TIPS

Accessibility in data visualization goes well beyond considerations for color blindness. This is something I personally understand much better after my interactions with Amy Cesal in preparation for bringing you the following post. Back in April, we posted a short article here with some accessibility considerations in visual design and Amy sent me a private note basically expressing, “but wait...there’s more!” I love when content posted on the blog starts conversations like this—as I continue to learn, it means there’s more great content to share with you. Just be aware that if you reach out like this, I may talk you into doing some writing and sharing: that’s what has happened here when I invited Amy to pen a guest post and she graciously accepted!

4. How much to visualize?

Am I showing too little?
Am I showing too much?
The U.S. Murder Rate (per 100,000 people)

FactCheck.org chart based on FBI data
The danger of aggregating data too much, and presenting just averages and other statistical summaries.

Most places are pretty safe, and have likely remained down here (these aren’t real data points).
The danger of aggregating data too much, and presenting just averages and other statistical summaries.
5. How to visualize it?
What encodings should I use?
How to make design choices?
Figure 2 - Main nationalities of arriving migrants – 2016

Greece
Figure 2 - Main nationalities of arriving migrants – 2016

Greece

- Syria 47%
- Afghanistan 24%
- Iraq 15%
- Pakistan 5%
- Algeria 1%
- Palestine 1%
- Iran 3%
- Others 4%
Books to make design choices

Draft available online: https://serialmentor.com/dataviz/
6. What visual style to use?

Not all visualizations need to be minimalist.

Not all visualizations need to be flashy and innovative, either.
Standard visualizations:
Appropriate for graphics we use all the time

Congressional Budget Office

Alberto Cairo • University of Miami • www.thefunctionalart.com • Twitter: @albertocairo
Fully customized visualizations:
Appropriate for one-time use when we want to have expressive impact

http://visualoop.com/28792/portfolio-of-the-week-jaime-serra

Alberto Cairo • University of Miami • www.thefunctionalart.com • Twitter: @albertocairo
7. What words to add?

“The purpose of visualization is insight, not pictures” (Ben Shneiderman) and “visualization alone often can’t provide that insight” (me)
"I and my colleagues here at the FT, we really do think one of the most valuable things we can do as data visualization practitioners is add this expert annotation layer."

**John Burn-Murdoch**

Financial Times


“Design secrets behind the FT’s best charts of the year”

[https://www.ft.com/content/4743ce96-e4bf-11e7-97e2-916d4fbac0da](https://www.ft.com/content/4743ce96-e4bf-11e7-97e2-916d4fbac0da)
The purpose of visualization isn’t visualization per se. It’s helping people make sense of the world through a combination of visuals and words, written or spoken.

Or, as Ben Shneiderman wrote, “the purpose of visualization is insight, not pictures.”