MAKING DECISIONS
IN VISUALIZATION

Alberto Cairo

Coda.Br 2020
We are living through a golden age of information graphics
NYTimes’ Dialect Quiz

Why outbreaks like coronavirus spread exponentially, and how to “flattenn the curve”

By Harry Stevens  March 14, 2020

PLEASE NOTE
The Washington Post is providing this story for free so that all readers have access to this important information about the coronavirus. For more free stories, sign up for our daily Coronavirus Updates newsletter.

https://www.washingtonpost.com/graphics/2020/world/corona-simulator/
The Post’s visual journalism, which involves staff throughout the newsroom, has attracted large audiences and contributed to record subscriber growth. Six of the seven most visited stories in The Washington Post’s history have been graphics, including the coronavirus simulator that became the most visited article in The Post’s history, with more than three times as many visits as the second. It also includes this year’s Democratic candidate quiz, which set the record for converting readers to subscribers.
At the same time, we’re still dealing with some misconceptions and myths:

1. “A picture is worth a thousand words”
2. “Visualization is intuitive”
3. “The data should speak for itself”
4. “Show, don’t tell!”
5. “Learn the rules of visualization (by reading this book by [insert author’s name here])!”
Visuals are often ambiguous
Visuals are often ambiguous
Visualizations can’t be designed based just on our personal preferences—although these are important.

**Visualization is a bit like 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.

Instead, when designing visualizations, we need to be guided by reasoned, justifiable choices.
“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
1. Why should my visualization exist?

Do the potential benefits of designing my visualization outweigh the possible harm it might cause?
<table>
<thead>
<tr>
<th>YEAR</th>
<th>TEMP</th>
<th>YEAR</th>
<th>SIGMA</th>
<th>SIGMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>0.0569</td>
<td>1000</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1001</td>
<td>-0.1241</td>
<td>1001</td>
<td>0.240347</td>
<td>0.480694</td>
</tr>
<tr>
<td>1002</td>
<td>-0.1208</td>
<td>1002</td>
<td>0.240347</td>
<td>0.480692</td>
</tr>
<tr>
<td>1003</td>
<td>-0.1801</td>
<td>1003</td>
<td>0.240347</td>
<td>0.480694</td>
</tr>
<tr>
<td>1004</td>
<td>-0.0711</td>
<td>1004</td>
<td>0.240347</td>
<td>0.480693</td>
</tr>
<tr>
<td>1005</td>
<td>-0.1334</td>
<td>1005</td>
<td>0.240346</td>
<td>0.480692</td>
</tr>
<tr>
<td>1006</td>
<td>-0.0644</td>
<td>1006</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1007</td>
<td>0.0042</td>
<td>1007</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1008</td>
<td>-0.1298</td>
<td>1008</td>
<td>0.240347</td>
<td>0.480693</td>
</tr>
<tr>
<td>1009</td>
<td>-0.0296</td>
<td>1009</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1010</td>
<td>0.1187</td>
<td>1010</td>
<td>0.240347</td>
<td>0.480693</td>
</tr>
<tr>
<td>1011</td>
<td>-0.1252</td>
<td>1011</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1012</td>
<td>-0.1634</td>
<td>1012</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1013</td>
<td>-0.0791</td>
<td>1013</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1014</td>
<td>-0.1120</td>
<td>1014</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1015</td>
<td>-0.1146</td>
<td>1015</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1016</td>
<td>-0.1206</td>
<td>1016</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1017</td>
<td>-0.0815</td>
<td>1017</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1018</td>
<td>-0.2031</td>
<td>1018</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1019</td>
<td>0.0305</td>
<td>1019</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1020</td>
<td>0.1098</td>
<td>1020</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1021</td>
<td>0.0244</td>
<td>1021</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1022</td>
<td>-0.0743</td>
<td>1022</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1023</td>
<td>-0.0323</td>
<td>1023</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1024</td>
<td>-0.0434</td>
<td>1024</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
</tbody>
</table>

Detecting patterns
Detecting patterns

Michael E. Mann, Raymond S. Bradley, and Malcolm K. Hughes
Intergovernmental Panel on Climate Change (IPCC), Third Report, 2001

Alberto Cairo • University of Miami • www.thefunctionalart.com • Twitter: @albertocairo
"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
Published Dec. 23, 2012 (the Sandy Hook Elementary School shooting was on Dec. 14)

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?
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?
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.
El no a la independencia de Cataluña gana al sí por primera vez desde 2012

El ‘CIS catalán’ constata que el apoyo a la secesión cae un 9% en los últimos dos meses

PERE RÍOS
Barcelona - 19 DIC 2014 - 15:51 EST

https://elpais.com/ccaa/2014/12/19/catalunya/1418984873_128596.html
For the first time since Catalan leader Artur Mas began his ongoing independence drive in 2012, a survey shows that a majority in the region would reject secession if a referendum were held now.

The latest poll by the Catalan executive’s Opinion Studies Center (CEO) shows that 45.3 percent of citizens would vote no to the question: “Would you like Catalonia to become an independent state?” compared with 44.5 percent who would support the move.

*Do you want Catalonia to become an independent state?*

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>45.3%</td>
</tr>
<tr>
<td>Yes</td>
<td>44.5%</td>
</tr>
<tr>
<td>No answer</td>
<td>10.2%</td>
</tr>
</tbody>
</table>
For the first time since Catalan leader Artur Mas began his ongoing independence drive in 2012, a survey shows that a majority in the region would reject secession if a referendum were held now.

The latest poll by the Catalan executive’s Opinion Studies Center (CEO) shows that 45.3 percent of citizens would vote no to the question: “Would you like Catalonia to become an independent state?” compared with 44.5 percent who would support the move.

Do you want Catalonia to become an independent state?

<table>
<thead>
<tr>
<th>Option</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>45.3%</td>
</tr>
<tr>
<td>Yes</td>
<td>44.5%</td>
</tr>
<tr>
<td>No answer</td>
<td>10.2%</td>
</tr>
</tbody>
</table>

Margin of error: +/-2.95 at 95% confidence level

“The margin of error of the poll is 2.95, a relevant fact considering the tight difference between the YES and the NO to the independence of Catalonia”
For the first time since Catalan leader Artur Mas began his ongoing independence drive in 2012, a survey shows that a majority in the region would reject secession if a referendum were held now.

The latest poll by the Catalan executive’s Opinion Studies Center (CEO) shows that 45.3 percent of citizens would vote no to the question: “Would you like Catalonia to become an independent state?” compared with 44.5 percent who would support the move.

<table>
<thead>
<tr>
<th>Do you want Catalonia to become an independent state?</th>
<th>Do you want Catalonia to become an independent state?</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>45.3%</td>
<td>45.3%</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>44.5%</td>
<td>44.5%</td>
</tr>
<tr>
<td>No answer</td>
<td>No answer</td>
</tr>
<tr>
<td>10.2%</td>
<td>10.2%</td>
</tr>
</tbody>
</table>

Margin of error: +/-2.95 at 95% confidence level

“The margin of error of the poll is 2.95, a relevant fact considering the tight difference between the YES and the NO to the independence of Catalonia”
Disclosing limitations and uncertainty

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

Collection of papers about visualizing uncertainty:
https://www.dropbox.com/sh/jk4ginyai6ylqu/AABvqdyTIhltyFN9nKNHyX9Ba?dl=0

https://graphics.cs.wisc.edu/Papers/2014/CG14/Preprint.pdf
3. How much to visualize?

Am I showing too little?
Am I showing too much?
The danger of aggregating data too much, and presenting just averages and other statistical summaries.
The danger of aggregating data too much, and presenting just averages and other statistical summaries.

Some places are so far up that they skew the national rate.

Most places are pretty safe, and have likely remained down here (these aren’t real data points).

FactCheck.org chart based on FBI data.
4. How to visualize it?
What types of charts or maps should I use?
What is the best way to organize the visualization?
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%
- Others 4%
- Iran 3%
- Palestine 1%
- Algeria 1%
### Visual vocabulary

**Designing with data**

There are many ways to visualize data. How do we know which one is right? How do we categorize them? The way we decide which data visualization is right may be different for different purposes. For some visualizations, we may want to have more information. For some, we may want to have a more intuitive and meaningful data visualization.

**ft.com/vocabulary**


Good journalism isn't just showing charts. It's also about explaining them: twitter.com/brianstelter/s...

This infographic is worth a thousand words – showing why "social distancing" and other protective measures helps to slow an outbreak. Hat tips to CDC, @theeconomist, @drewaharris, and @CT_Bergstrom

https://twitter.com/AlbertoCairo/status/1236773377865658370
“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
Every outcome in our simulations

All possible Electoral College outcomes for each candidate, with higher bars showing outcomes that appeared more often in our 40,000 simulations

https://projects.fivethirtyeight.com/2020-election-forecast/
5. What 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


Alberto Cairo • University of Miami • www.thefunctionalart.com • Twitter: @albertocairo
Fully customized style:
Appropriate for one-time use when we want to provoke curiosity, surprise —or simply a smile

https://jaimeserra-archivos.blogspot.com/
Fully customized style:
Appropriate for one-time use when we want to provoke curiosity, surprise —or simply a smile
E se todos os mortos por Covid-19 no Brasil fossem seus vizinhos?

Descubra o que aconteceria com a sua vizinhança caso o epicentro da epidemia de Covid-19 no Brasil fosse sua casa.

Publicado em 24 de julho de 2020.
Dados atualizados em 1 de novembro de 2020.

The purpose of visualization isn’t visualization per se. The purpose of visualization is to help people make sense of the world through a combination of visuals and words.
The End.

www.thefunctionalart.com, www.albertocairo.com, alberto.cairo@gmail.com