BEST PRACTICES
for Data Visualization

CSTE Public Health Drug Overdose Surveillance
Training Series for Local/Territorial Jurisdictions

CSTE

March 2021

Humans comprehend by visualization.

New studies in what has become known as the architecture of data visualization provides insight into the techniques of graphical representation of data, allowing for ease of comprehension and analysis. Choosing the proper visualization, graph or chart impacts the desired results and works to serve a specific purpose. Effective visualization makes complex data more accessible, understandable, and persuasive. This article offers best practices for data visualization.

This work is supported in part by Cooperative Agreement Number 1 NU38OT0002927-021-00 awarded to the Council of State and Territorial Epidemiologists from the Centers for Disease Control and Prevention (CDC). Content developed by Somya Mohanty, PhD and Prashanti Manda, PhD. Support for this workbook came from Allyson Kelley and Associates PLLC. The views expressed do not necessarily reflect the official priorities of the Department of Health and Human Services nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S government.

Table of Contents

Best Practices for Data Visualization

Speak to a specific audience 4

Choose the right visual 4

Provide context in your visuals 8

Keep things simple and digestible 8

Practice graphical integrity 9

Maximizing the data-ink 10

Avoid chartjunk 11

Use classic design techniques with small multiples 11

Show trends of time for dramatic messages 12

Design Principles for Visualizations 13

Enclosure 13

Similarity 14

Proximity 14

Closure 15

Symmetry 15

Do's and Don'ts 16

Speak to a specific audience

Consider the audience and the message you are trying to convey. If you are you are trying to convey the effects of climate change to a group of legislators, your goal is to inspire urgency. In this case, highlight the drastic effects of climate change along with a temporal component to make the repercussions imminent. Tailor your visuals to convey the right message.

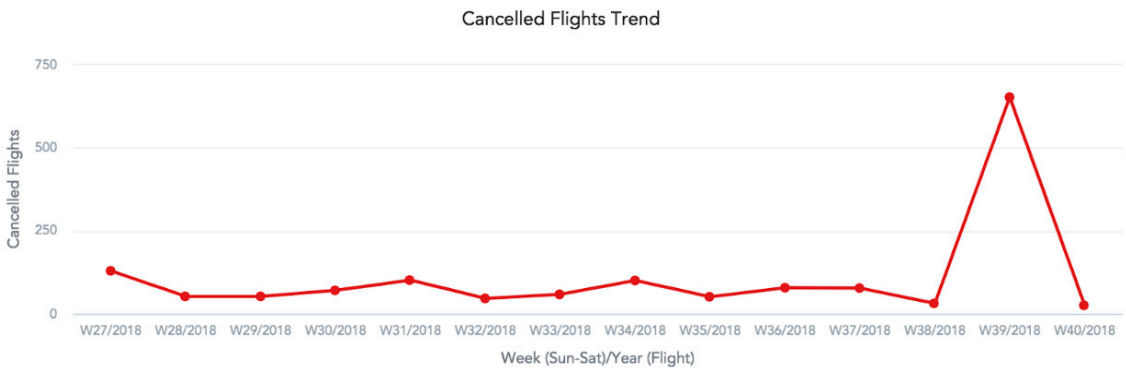
Choose the right visual

Tables show quantitative information but they tend to be overwhelming. Provide a narrative to explain the information.

Channel ▲	Campaign	Total Cost	MQLs	Cost per MQL
Content Syndication	E-Commerce Data Matt…	$19,345.26	40	$483.63
Digital	Game Time Savings	$21,035.77	90	$233.73
Digital	Many-to-Many Healing	$22,609.90	46	$491.52
Email	Analytics Maturity Cam…	$2,031.05	24	$84.63
Email	Quarterly Trends Camp…	$13,037.58	36	$362.16
Tradeshow	CIO Summit	$128,089.00	381	$336.19
Sum		**$206,148.56**	**617**	

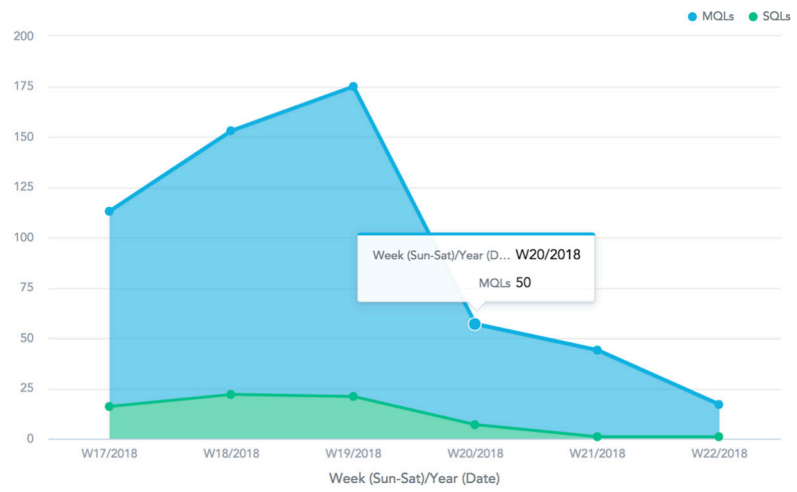
Source: https://www.gooddata.com/blog/5-data-visualization-best-practices-0

Use a line chart to track changes over time or to show the relationship between two or more variables.



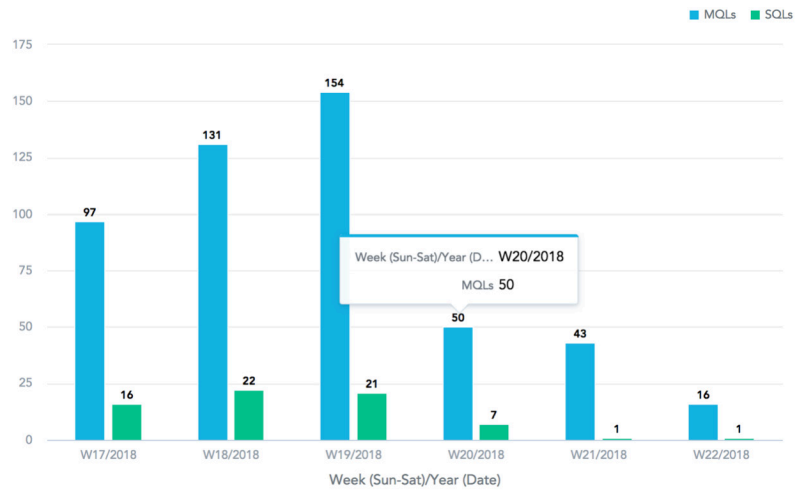
Source: https://www.gooddata.com/blog/5-data-visualization-best-practices-0

Use an area chart to highlight differences in the area covered under the lines.



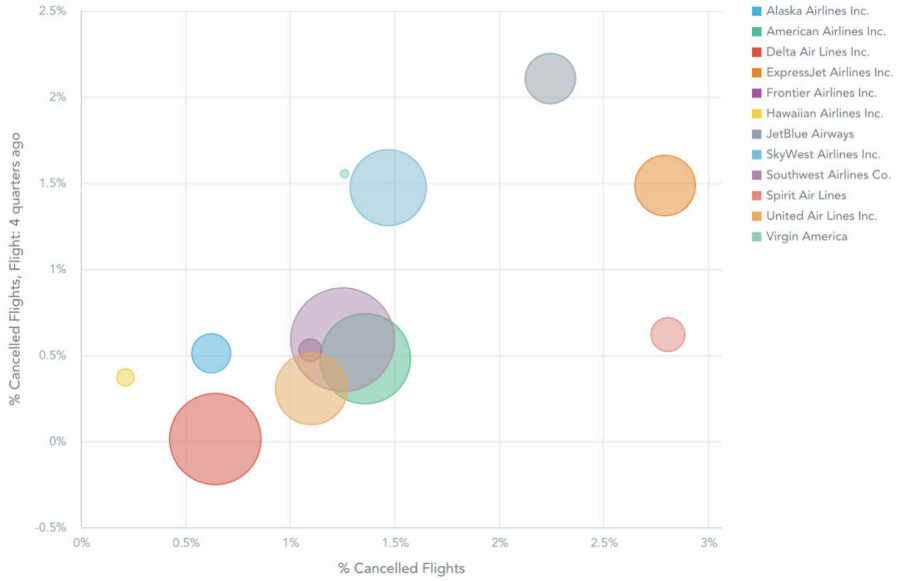
Source: https://www.gooddata.com/blog/5-data-visualization-best-practices-0

Use a bar chart to compare different quantities or to show differences among categories. Different colors show different categories.



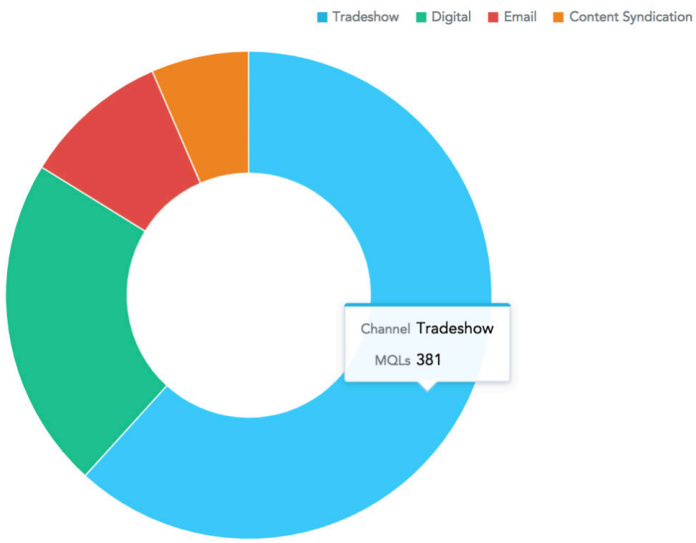
Source: https://www.gooddata.com/blog/5-data-visualization-best-practices-0

Use a bubble chart to add a third dimension. The size of each bubble represents an additional aspect.



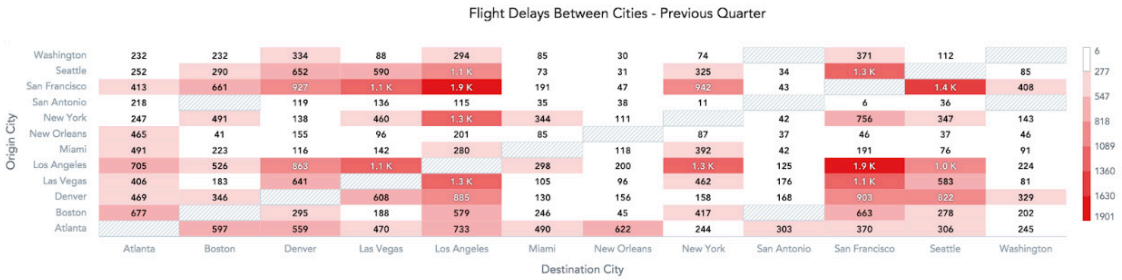
Source: https://www.gooddata.com/blog/5-data-visualization-best-practices-0

Pie charts compare parts of a whole. Donut cards are pie charts with a hole in the center. Avoid comparing two pie charts, it makes for visual confusion.



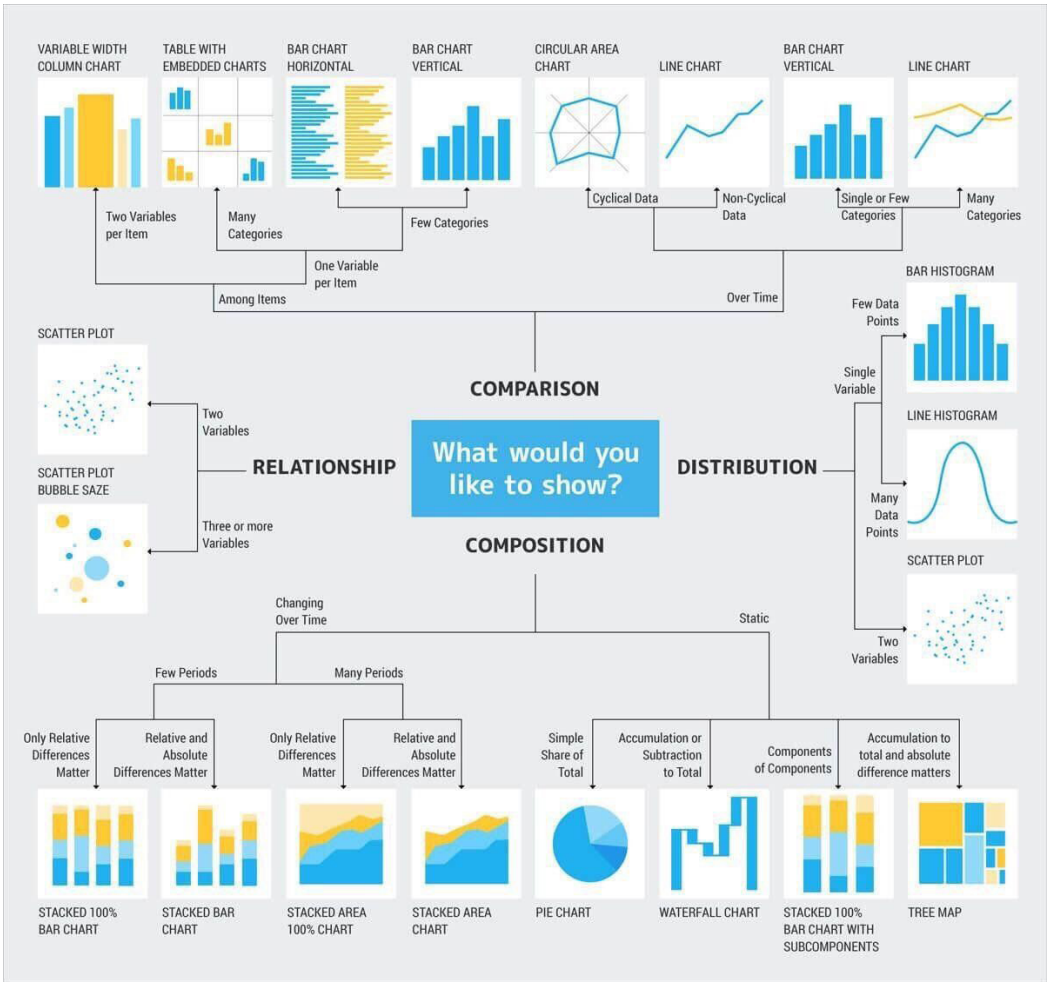
Source: https://www.gooddata.com/blog/5-data-visualization-best-practices-0

Heatmaps show an array of quantities in a color-coded fashion and show correlations between a set of variables.



Source: https://www.gooddata.com/blog/5-data-visualization-best-practices-0

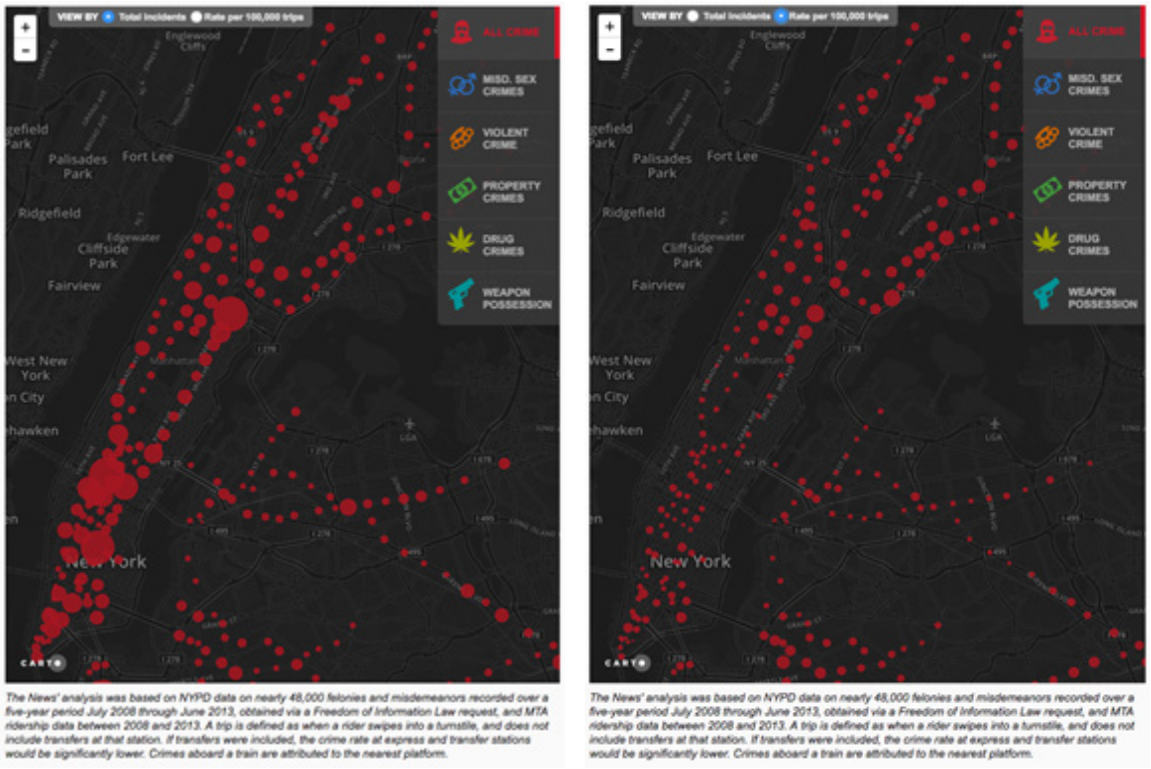
Use a comprehensive guideline, or cheat sheet, to identify your visualization applications.



Source:

Provide context in your visuals

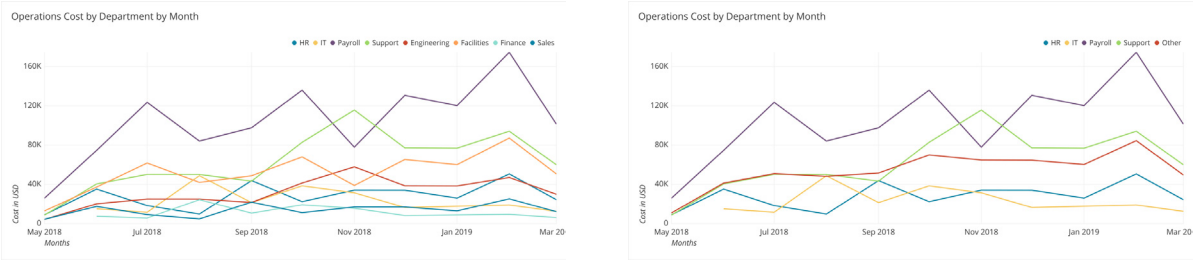
The left image shows number of crimes at different subway locations. The image on the right shows the same data as the number of crimes per 100,000 trips through a station. This plot shows a completely different story.



Source: https://www.nydailynews.com

Keep things simple and digestible

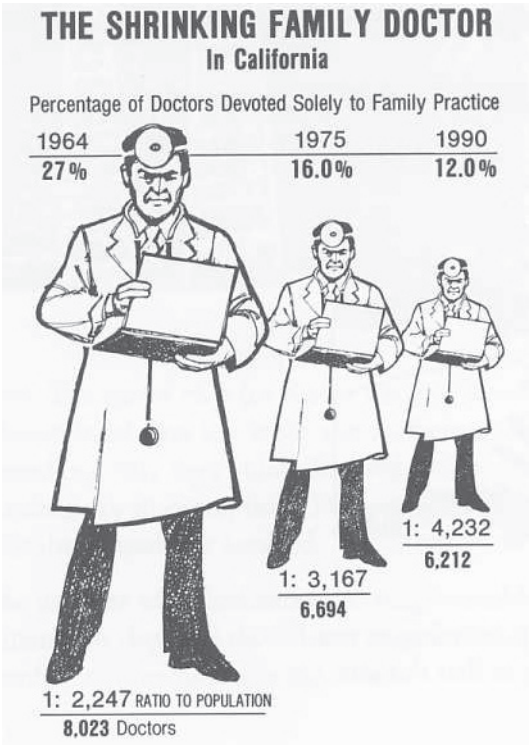
The first image shows too many categories and can easily overwhelm the reader. The differences in trends are ambiguous. The second plot reduces the number of categories to increase simplicity.



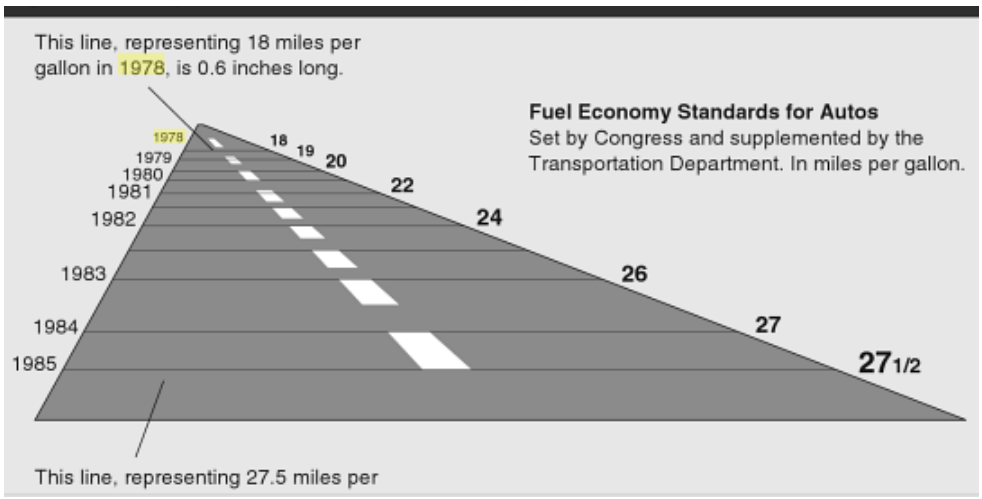
Source: https://chartio.com

Practice graphical integrity

In the following images, the change of proportion in the visuals is overblown as compared to changes in absolute numbers. This is misleading and is referred to as the "Lie Factor." In other words, the scale of the graphic should always correspond to changes in the data being represented.



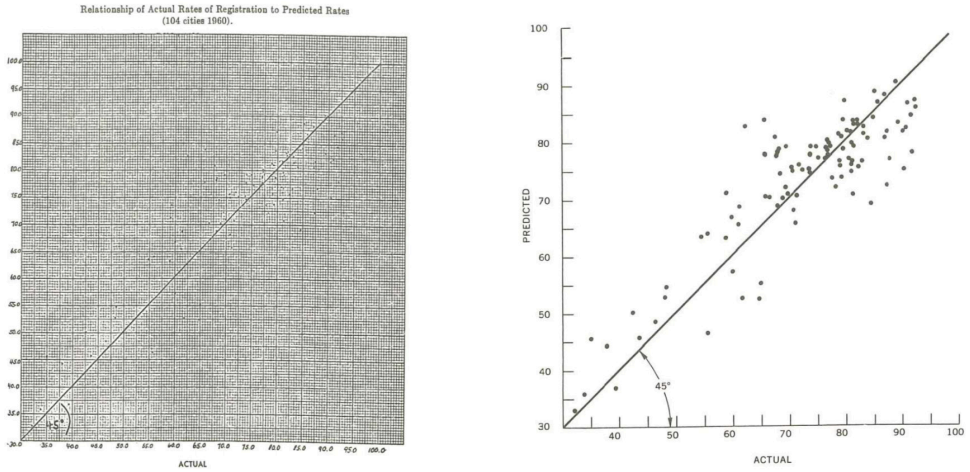
Source:



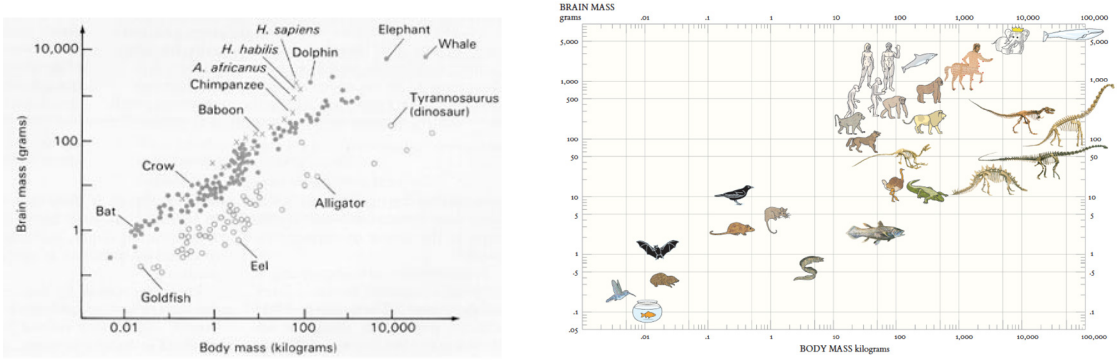
Source: Tufte, 1991

Maximizing the data-ink

Data-ink corresponds to the parts that represent data on a plot. Maximize data-ink and erase as much non-data-ink as possible. The left image is redesigned to be more intuitive and effective with the plot on the right.

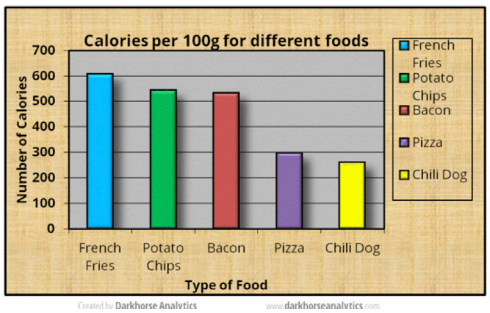


Source: Tufte, Edward R. 1986. The Visual Display of Quantitative Information. Cheshire, CT, USA: Graphics Press

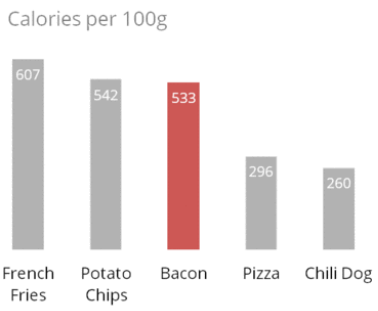


Source: Tufte, Edward R. Beautiful evidence. Vol. 1. Cheshire, CT: Graphics Press, 2006

BEFORE



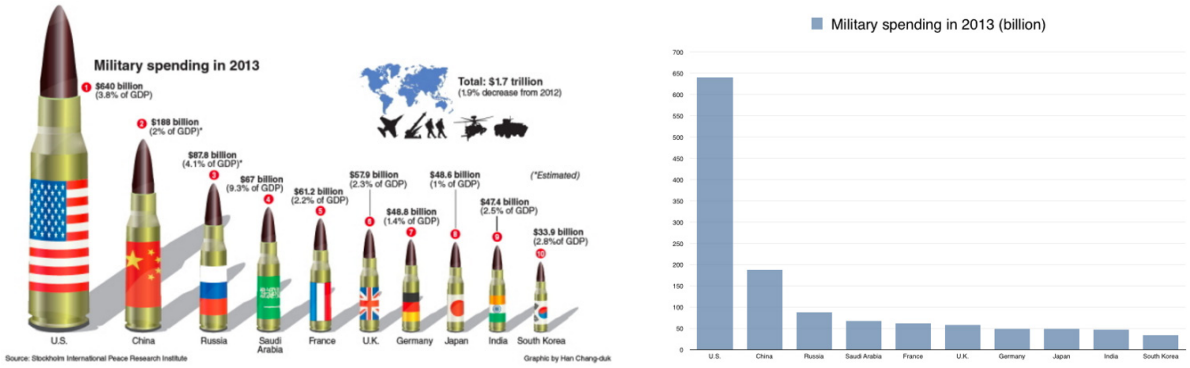
AFTER



Source: https://medium.com/@plotlygraphs/maximizing-the-data-ink-ratio-in-dashboards-and-slide-deck-7887f7c1fab

Avoid chartjunk

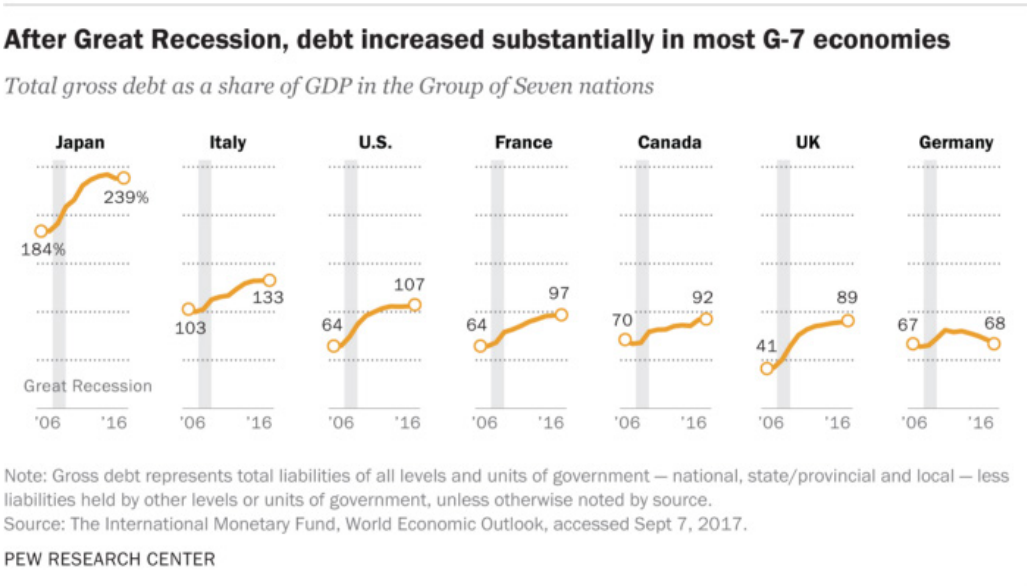
Avoid excessive and unnecessary use of graphics or elements that do not contribute directly to conveying the main message of the visualization.



Source: https://medium.com/@migle.rusteikaite

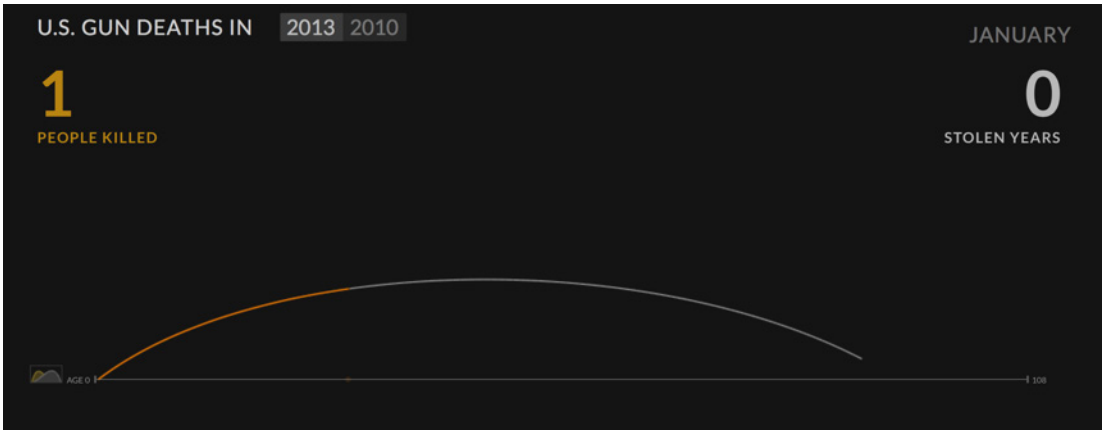
Use classic design techniques with small multiples

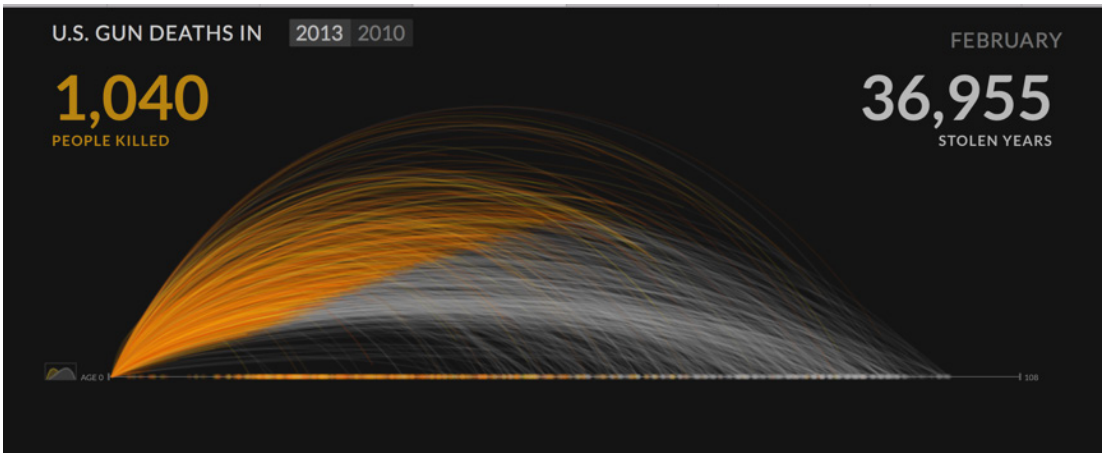
Display several charts in a consistent manner at a small size to show related trends.



Source:

Show trends of time for dramatic messages





Source: Periscope. 2018. "US Gun death in 2010 and 2013." https://guns.periscopic.com/?year=2013

DESIGN PRINCIPLES
for Visualizations

There are several principles of design put forth by several influential designers and visualization experts.

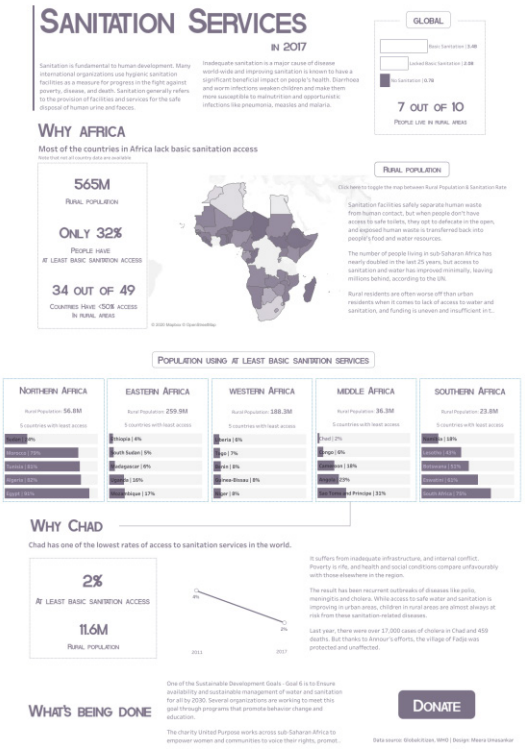
The schools of thought are:
1. Tufte's design principles
2. Melissa Anderson's principles
3. Gestalt's principles of design

Several of Tufte's principles have been incorporated into the best practices discussed above. Below, we discuss Gestalt's principles of design:

These principles are centered around seven pillars: enclosure, similarity, continuity, closure, connection, proximity and symmetry.

Enclosure

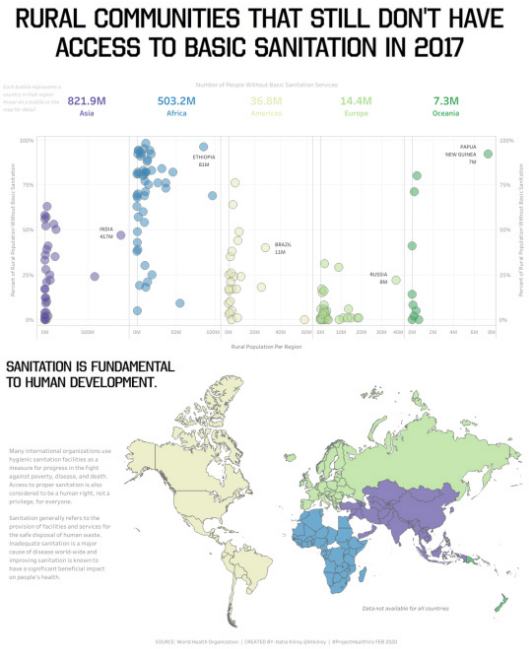
Objects placed within enclosures or boundary-like structures are perceived as a group. Use enclosures to cluster elements that are part of a group. In the below infographic, boxes are used to group key pieces of information.



Source: Meera Umasankar, Tableau representative

Similarity

When objects share similar features/attributes, such as shapes, colors, sizes, they are perceived to be in the same group. The visuals below use color, shapes, and sizes to identify similar groups.



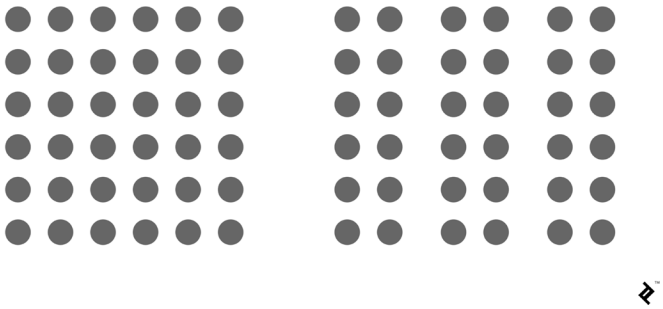
Source: Katie Kilroy https://public.tableau.com/profile/katie.kilroy

Proximity

Proximity of data elements in a plot signals that the elements belong to a group. Proximity is powerful enough to override differences in shapes/colors/etc.

Source: Andy Rutledge https://www.usertesting.com/blog/gestalt-principles
Source: Steven Bradley https://www.usertesting.com/blog/gestalt-principles

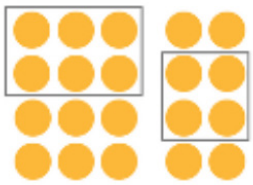
Proximity refers to how close elements are to each other. For example, the group of lines on the left is seen as one group, whereas; the group on the right is perceived as three separate groups. The only difference between the two groups of lines is the proximity of the lines.



https://www.toptal.com/designers/ui/gestalt-principles-of-design

Closure

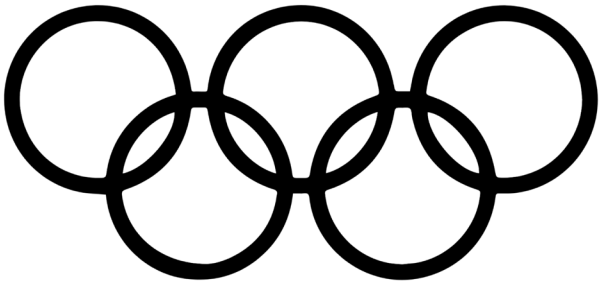
Enclose items that belong in a group.



Source:

Symmetry

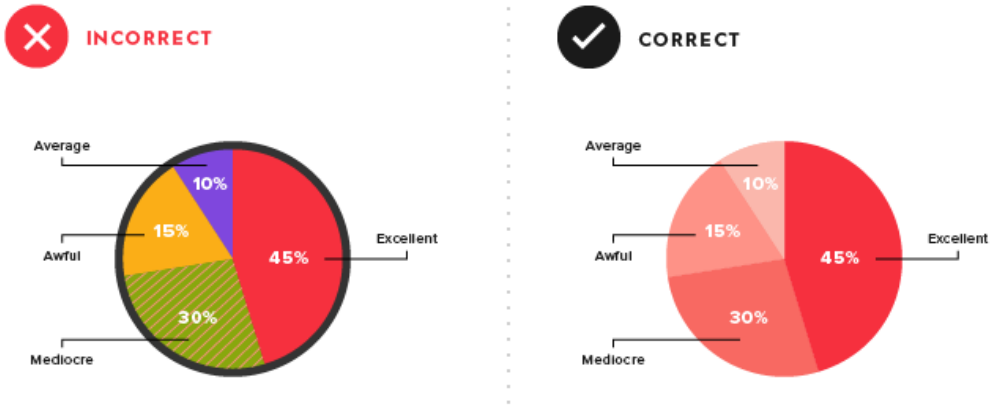
This principle says that your brain will perceive ambiguous shapes in as simple a manner as possible. For example, a monochrome version of the Olympic logo is seen as a series of overlapping circles rather than a collection of curved lines.



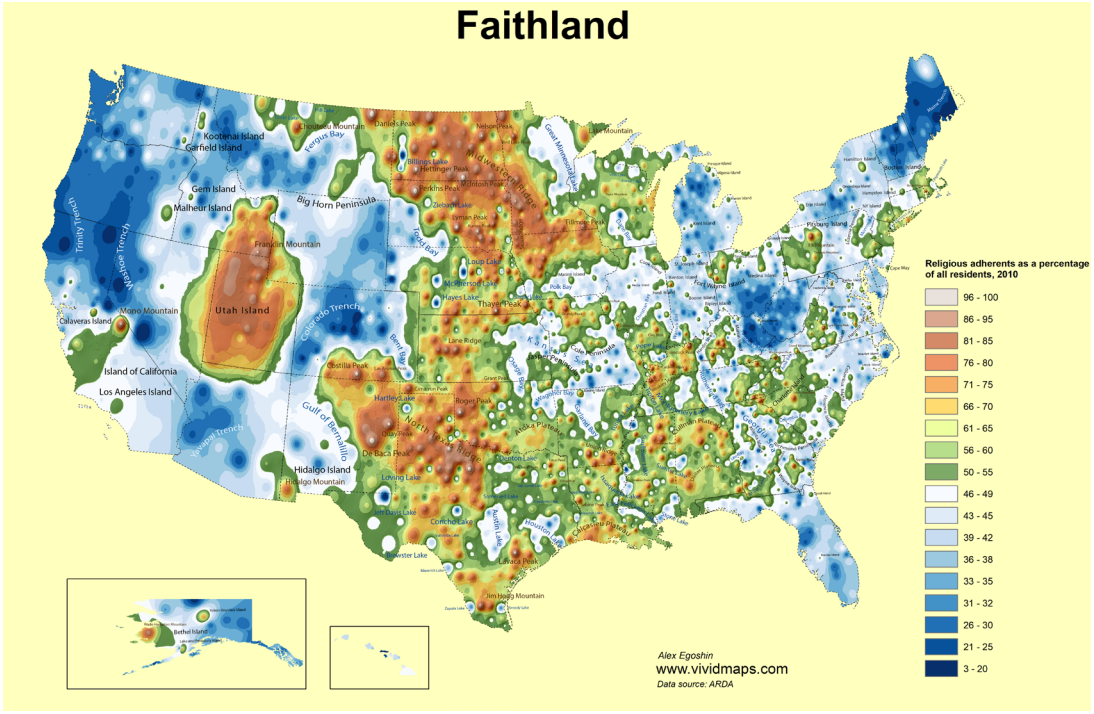
https://www.toptal.com/designers/ui/gestalt-principles-of-design

DO'S AND DON'TS

Don't use more than 6 colors in a single layout

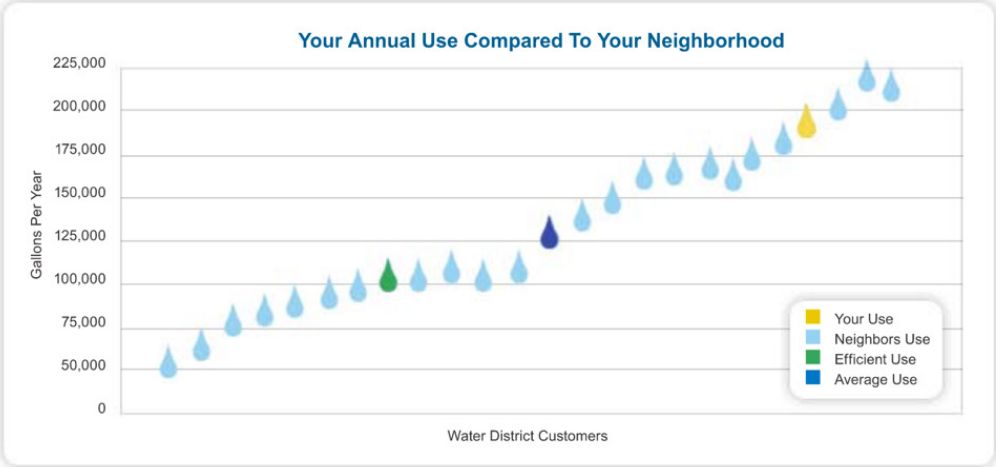


Source: https://www.columnfivemedia.com



Source: Visualization by Alex Egoshin at Vivid Maps

The colors allotted for the different categories are indistinguishable lending to poor understandability.

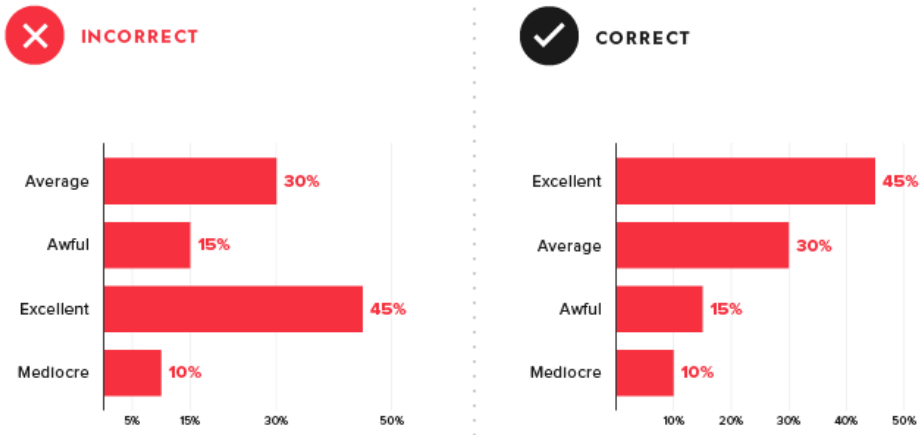


Source: Graph via Radian Developers

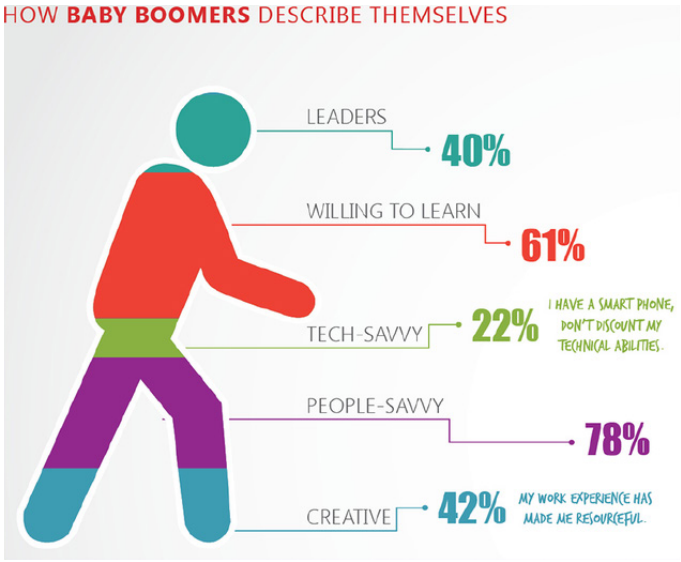
In contrast, the above graph is clearly defined, simple, and uses different colors to convey the message.

Use good axis measurements and labels

Use natural increments on your axes. Don't misrepresent your data by starting axes at arbitrary points. Arbitrary starting points exaggerate differences in data.
Always order your data for ease of understanding.



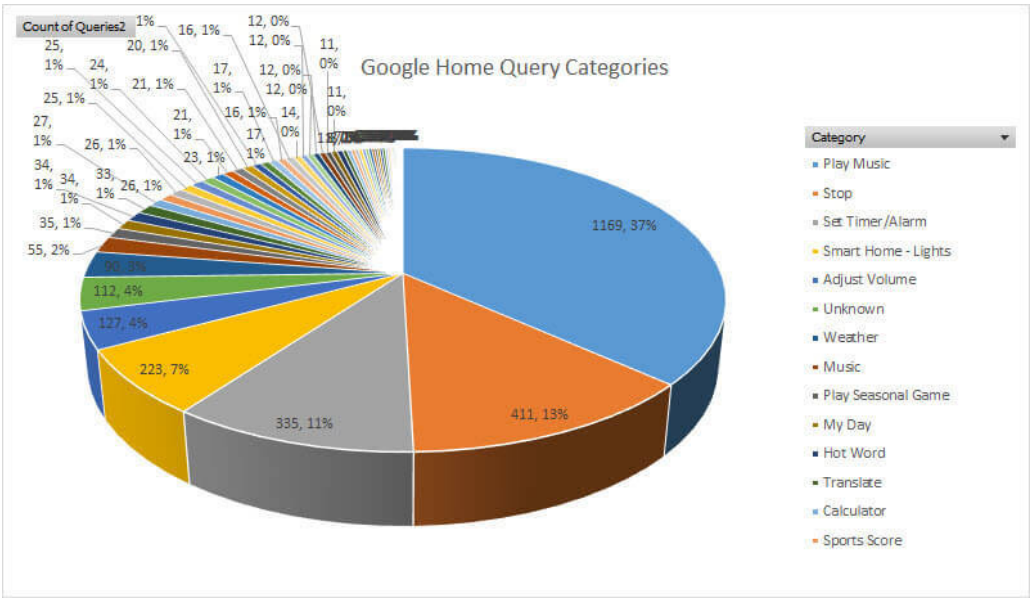
Source: https://www.columnfivemedia.com



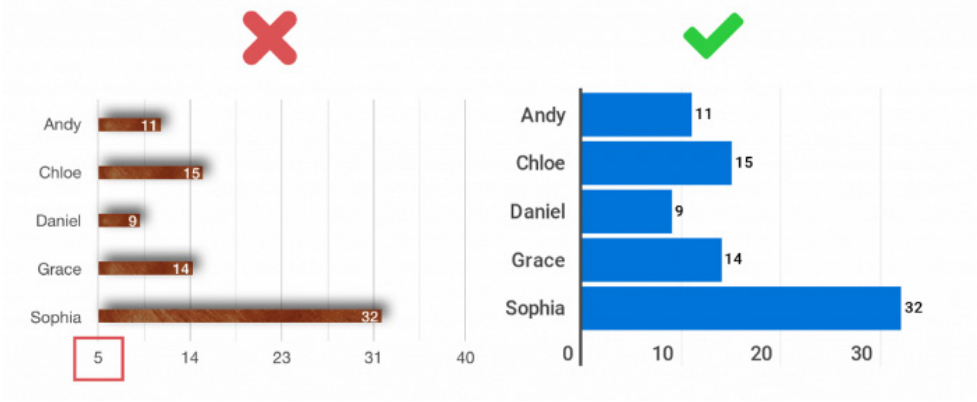
Source: https://www.columnfivemedia.com

The above graphic is misleading because the amount of visual space used for each category is not proportional to the corresponding percentage of the category. Also, some colors draw more attention than others.

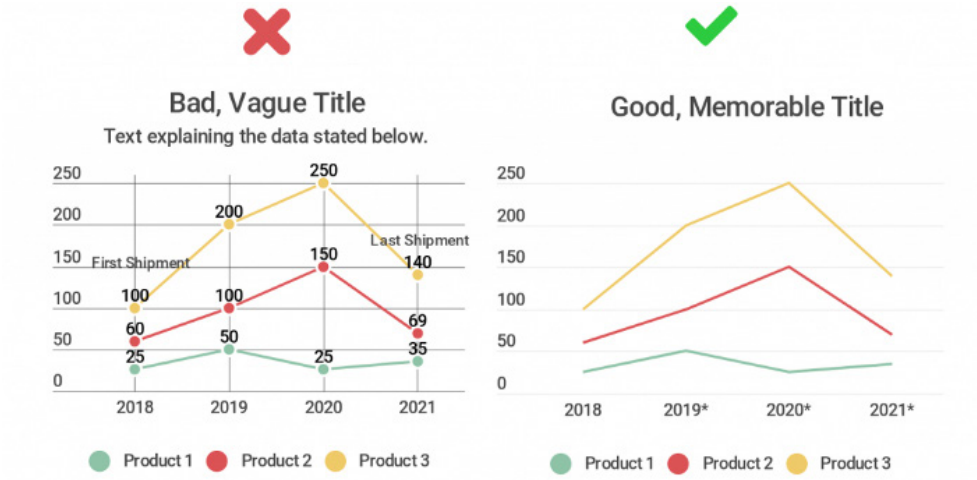
Don't present too much information



Source: WTF Visualizations



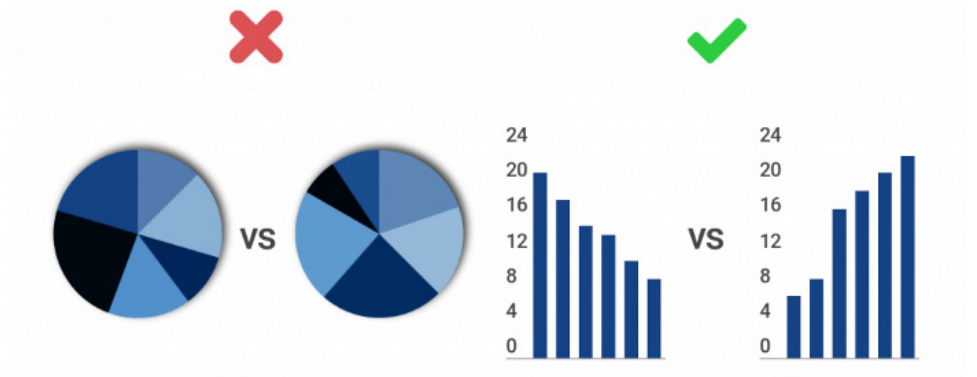
Source: https://medium.com/@Infogram



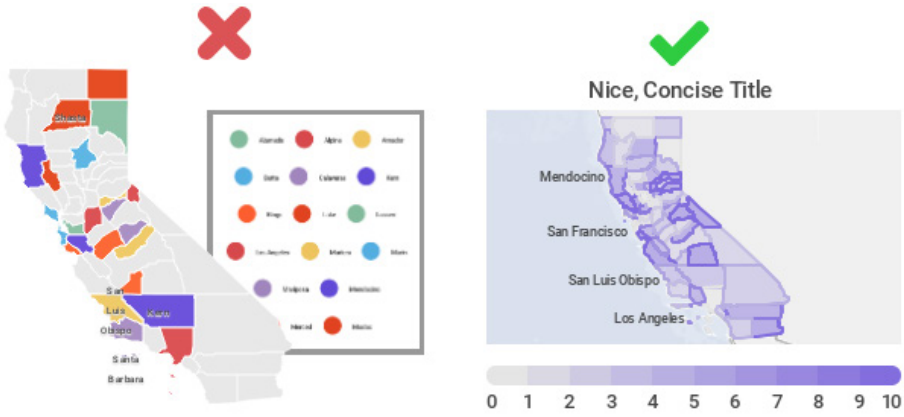
Source: https://medium.com/@Infogram

Don't overload the charts with labels, legends, and extra information

Be selective about the text you add to the chart. Make sure that the axes are labeled properly. Feel free to add a pop of color to highlight a data point that stands out.

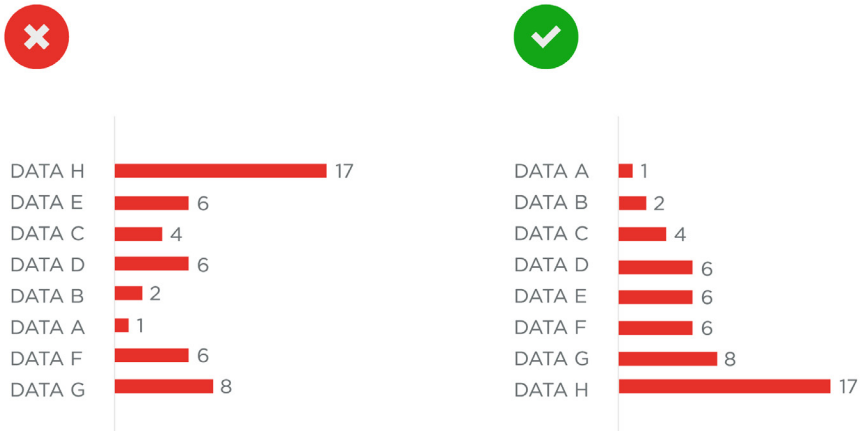


Source: https://medium.com/@Infogram



Source: https://medium.com/@Infogram

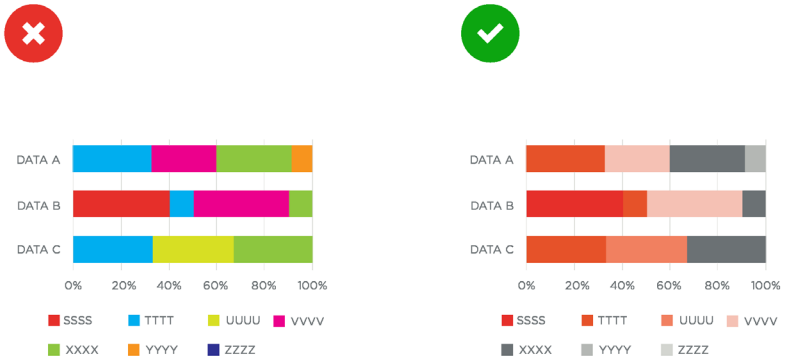
Order your data



Source: https://www.merrittgrp.com

Keep Your Colors to a Minimum

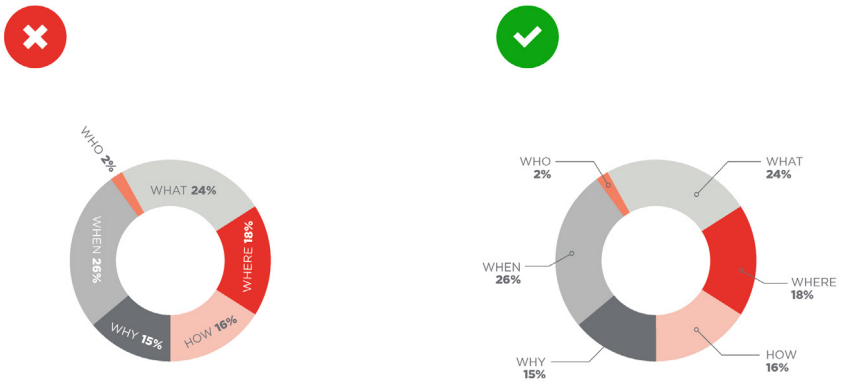
Too many colors can overwhelm a reader. A good rule of thumb is to use no more than 6 colors in one visual.



Source: https://www.merrittgrp.com

Label All Visual Elements Clearly

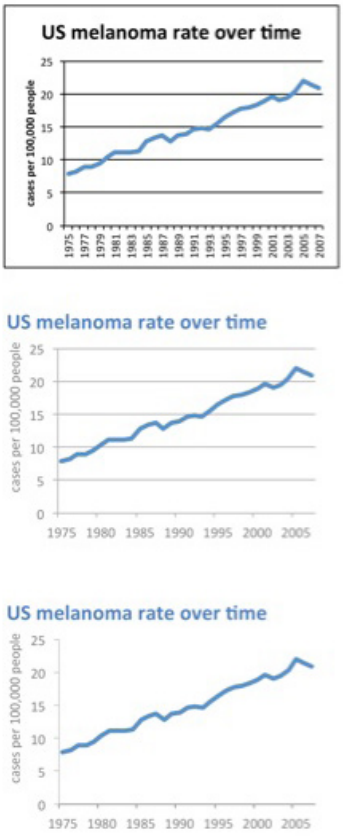
Keep labels easy to read and short. Make sure all visual elements are labeled.



Source: https://www.merrittgrp.com

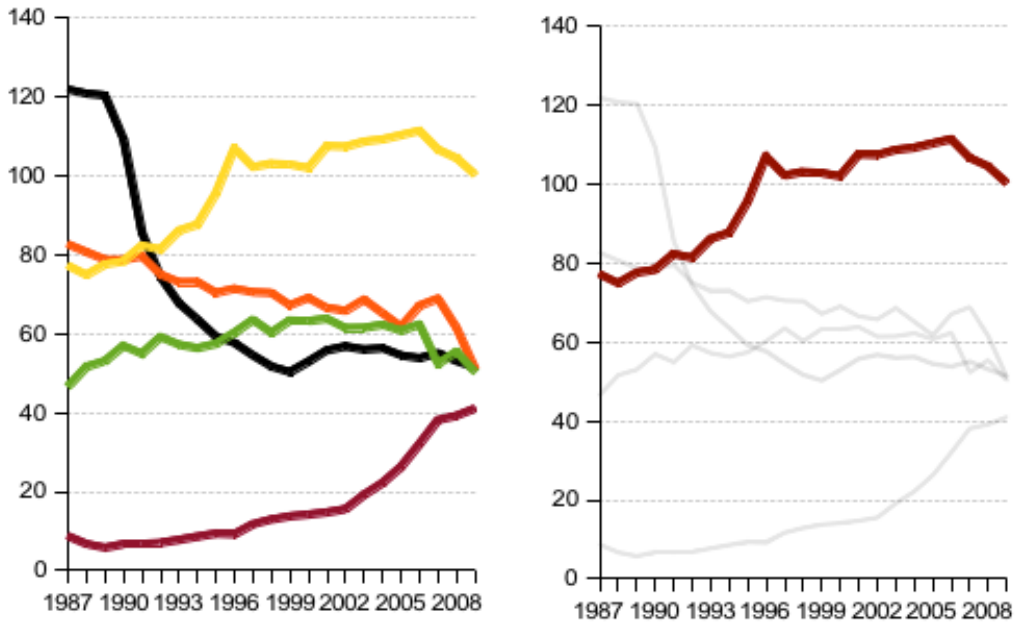
Simplify less important information

Eliminate distracting elements such as gridlines. Avoid rotating tick labels.



Source: https://guides.library.duke.edu/datavis/topten

Highlight the most important elements of your visual with color



Source:

Data visualization is an effective tool that achieves the goal of communicating information clearly and effectively.

Following these basic principles of design helps the reader to perceive objects easily and enables decision makers to view, understand and analyze complex concepts.

Supported by the Council of State and Territorial Epidemiologists
2635 Century Parkway NE
Suite 700
Atlanta, GA 30345
Tel: 770.458.3811

AKA PLLC
69705 Lake Drive
Sisters, OR 97759
Tel: 919.265.7549

CSTE