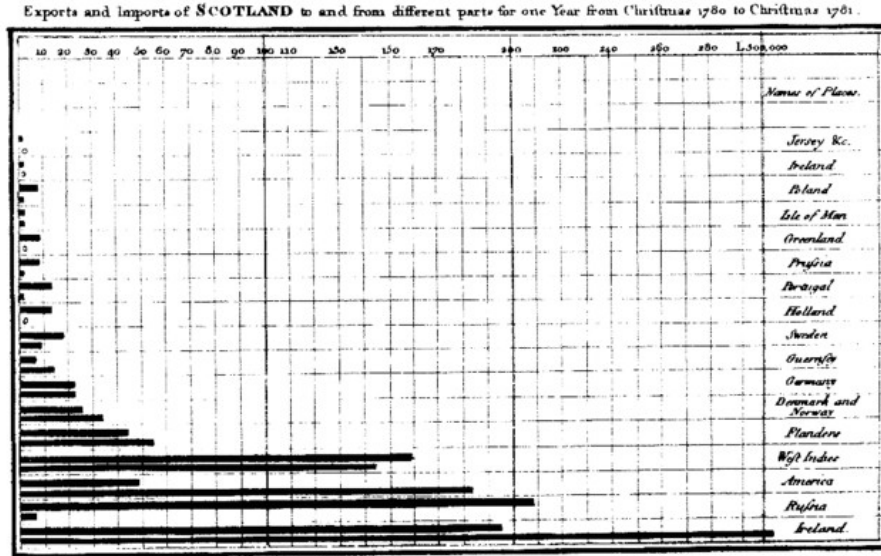


Dataviz

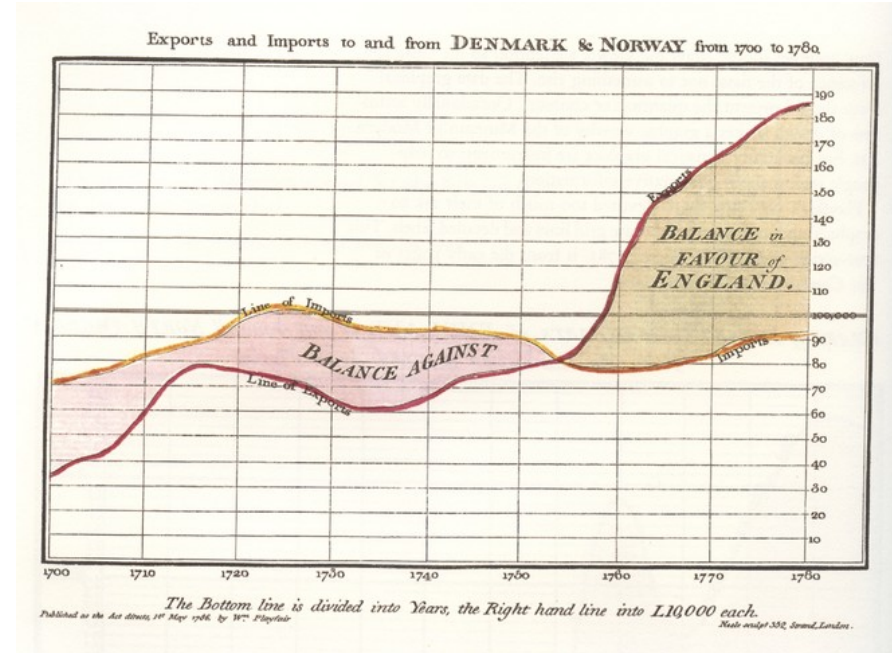


DATAVIZ: A BIT OF HISTORY

The pioneer: William Playfair
(1759 - 1823)

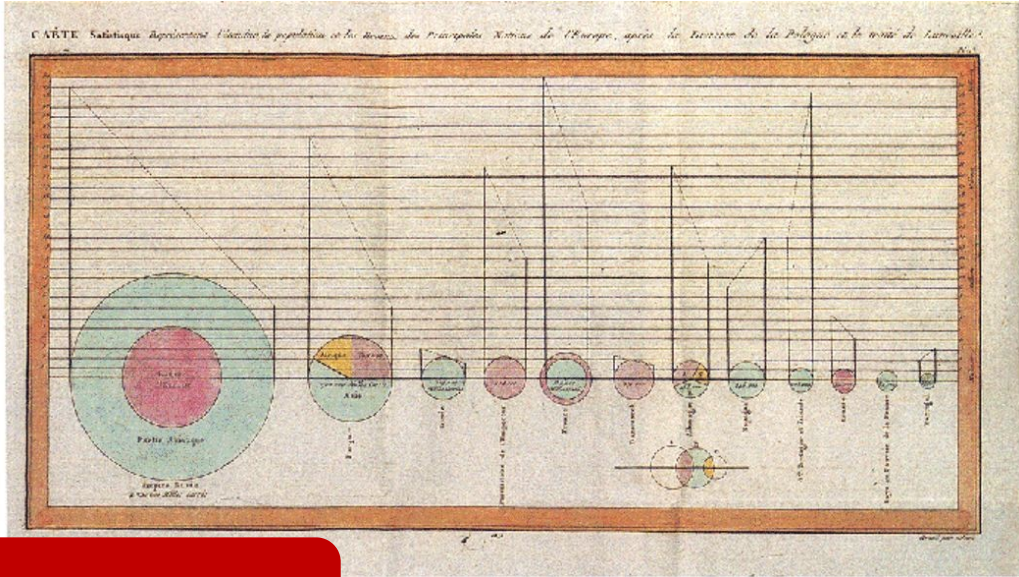
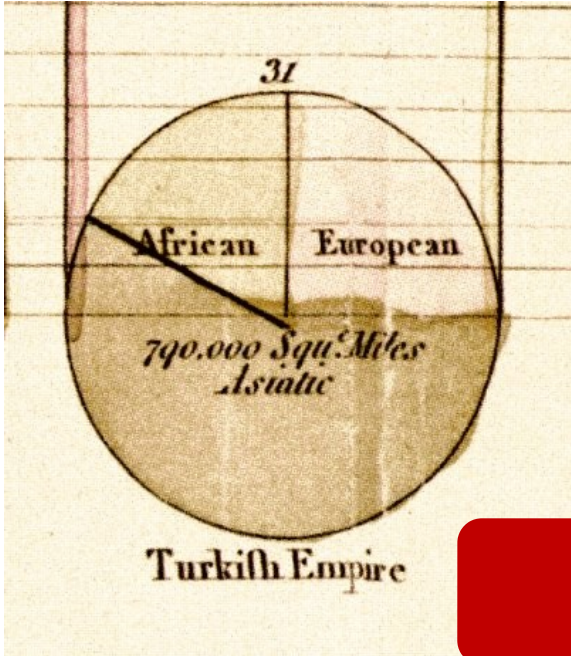


The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Red Lines Imports.
Published in the Art Magazine, June 7th 1786 by W^m Playfair
High street 152 & Strand, London.



Time series and Bar
chart (1786)

The pioneer: William Playfair
(1759 - 1823)

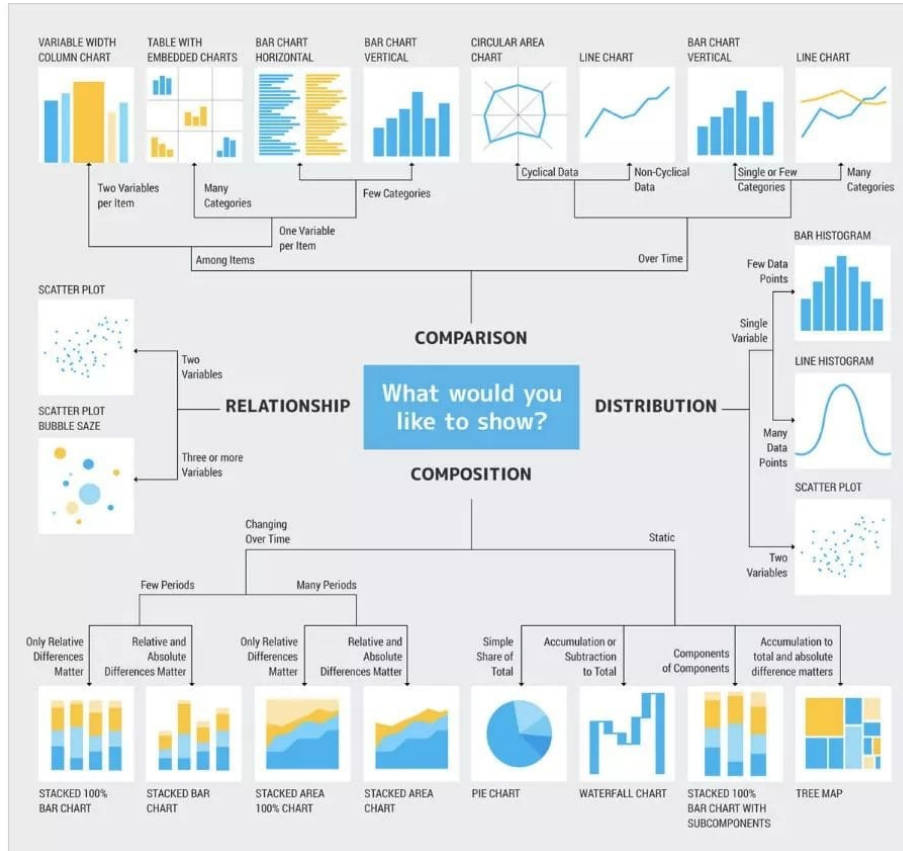


Pie chart (1801)

DATAVIZ: WHY & HOW?

- To transform raw data into relevant information
- To pick up only relevant information
- To give a message to an audience

DATAVIZ: HOW TO CHOOSE?



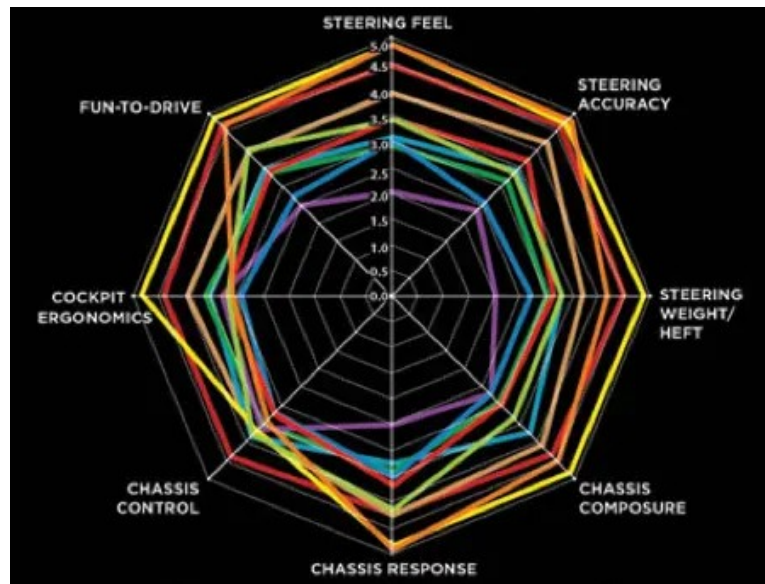
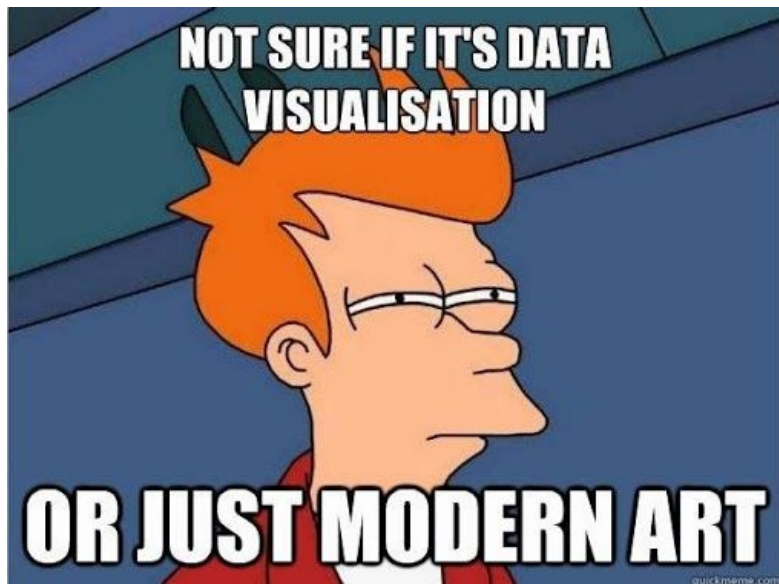
Best choice:

<https://www.data-to-viz.com/>

Good practices

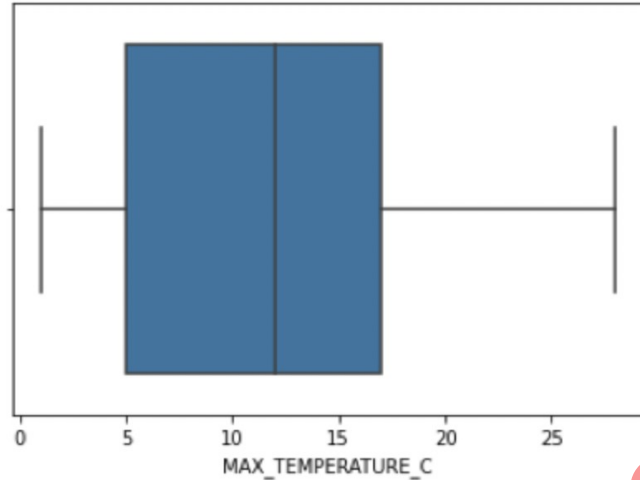
THE PITFALLS OF DATA VISUALISATION

STAY READABLE

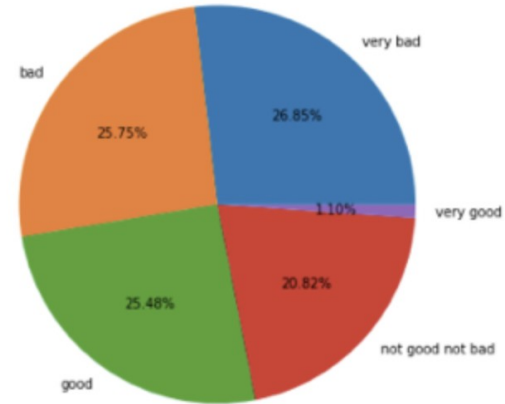


DEFINE YOUR AUDIENCE

Statistician
oriented



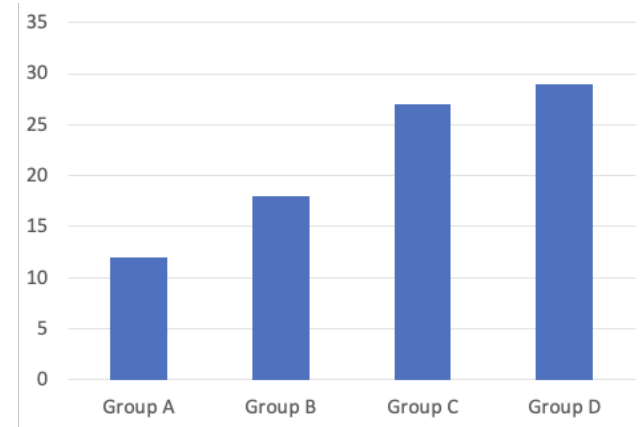
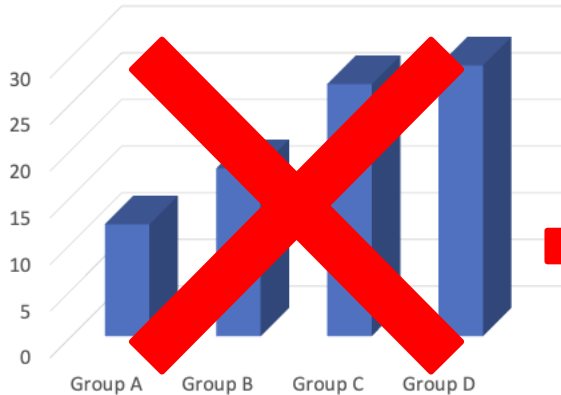
Communication
oriented



A dataviz is not exhaustive. It is a subjective (but ethical) choice.

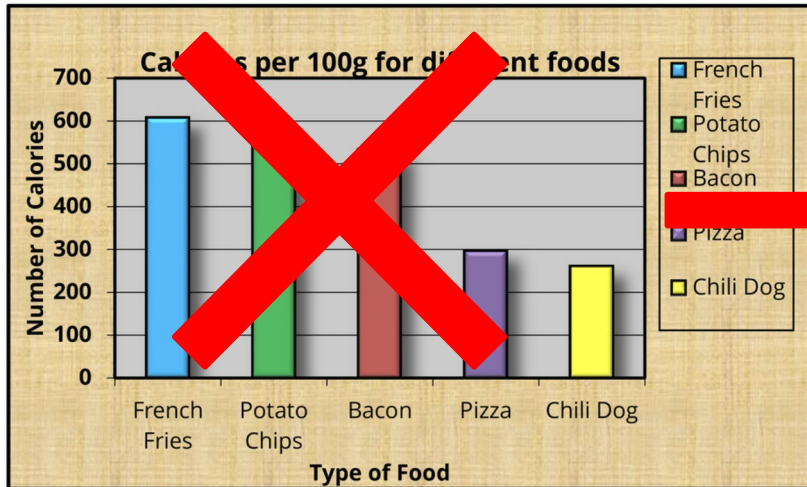
AVOID OVERLOADING THE GRAPHS

The simpler it is, the more readable it is. Also avoid all 3D effects, it usually doesn't bring much.

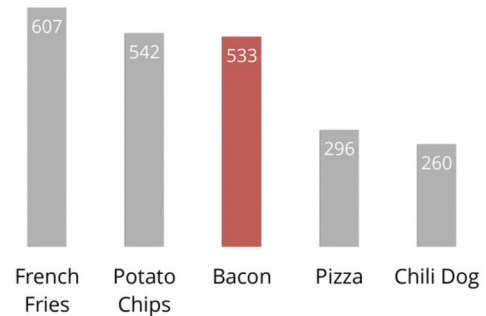


DATA-INK RATIO

Remove
to improve
(the **data-ink** ratio)

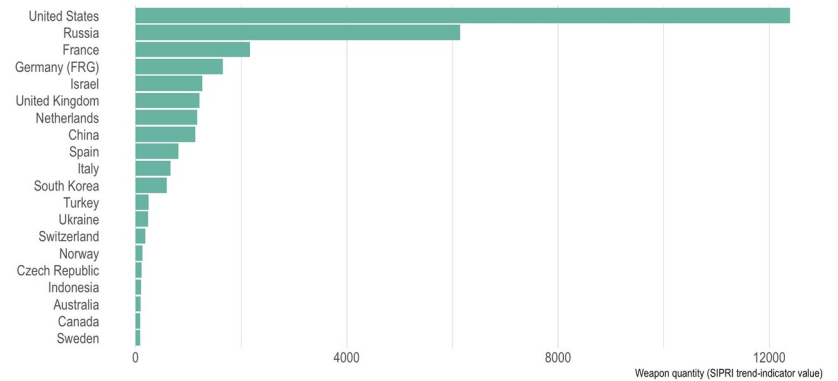
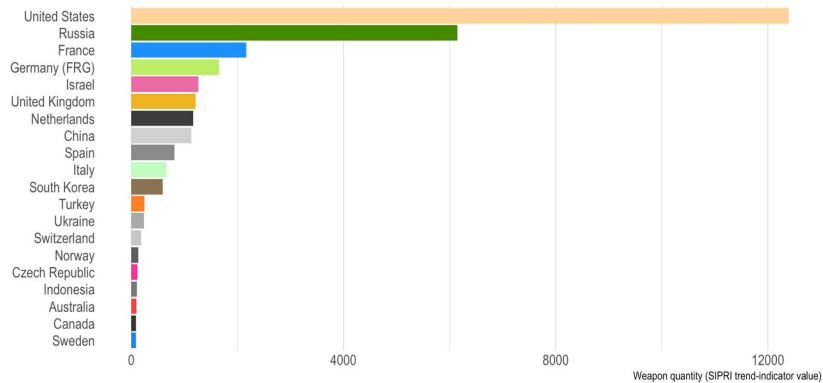


Calories per 100g



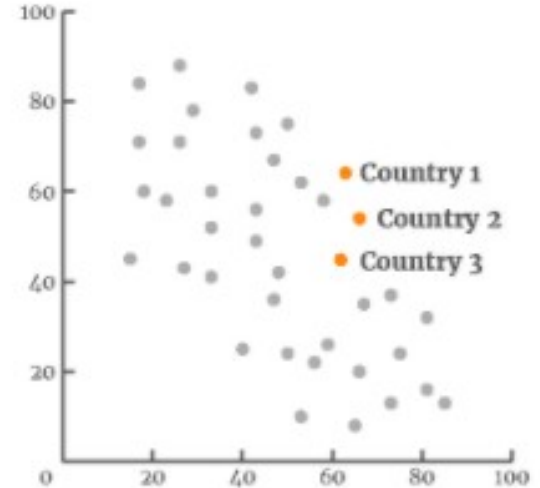
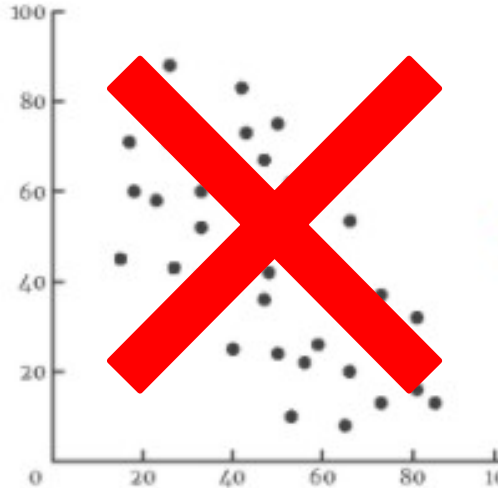
THE CHOICE OF COLORS

If colours are unnecessary, don't use them.



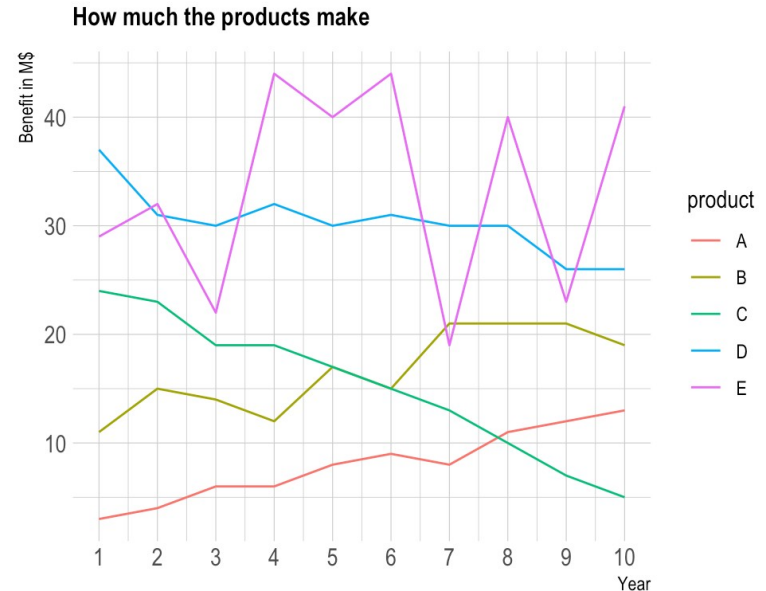
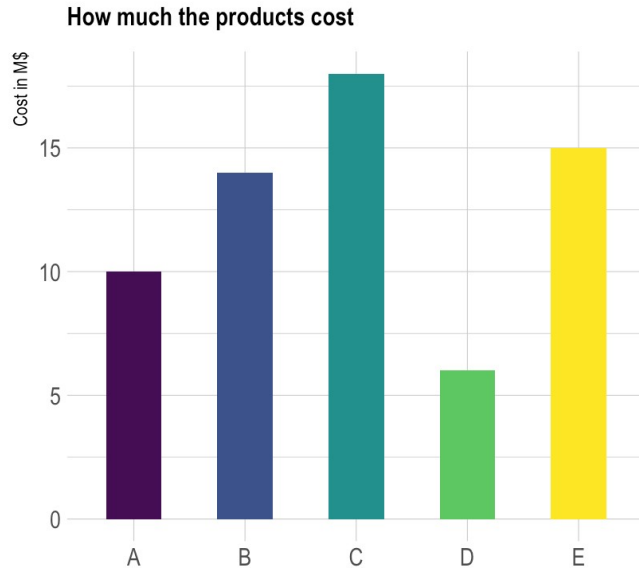
THE CHOICE OF COLORS

Don't hesitate to emphasize important information. The goal is to guide the reading to be as intuitive as possible.



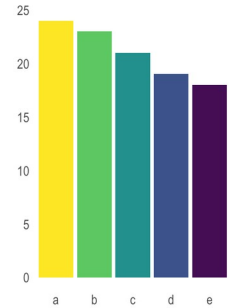
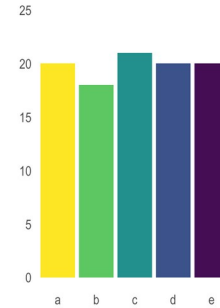
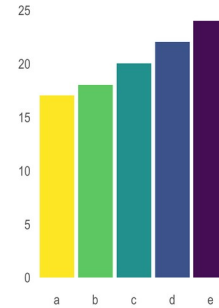
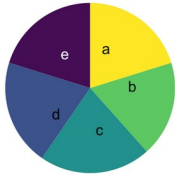
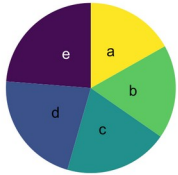
THE CHOICE OF COLORS

Always use the same colour for related graphics



THE MISUSE OF PIE CHARTS

The human eye is bad at quantifying angles...

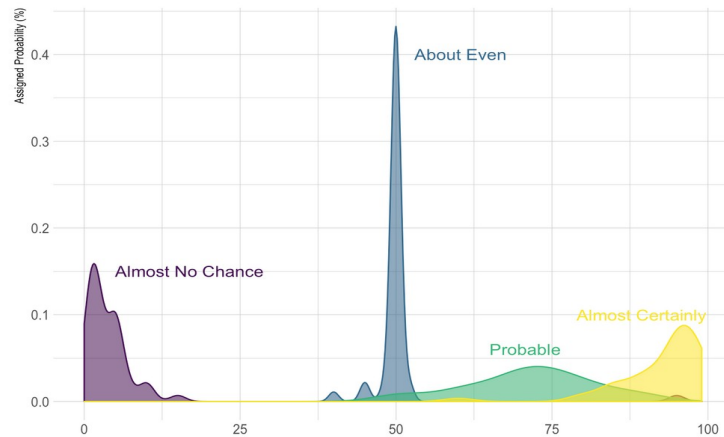
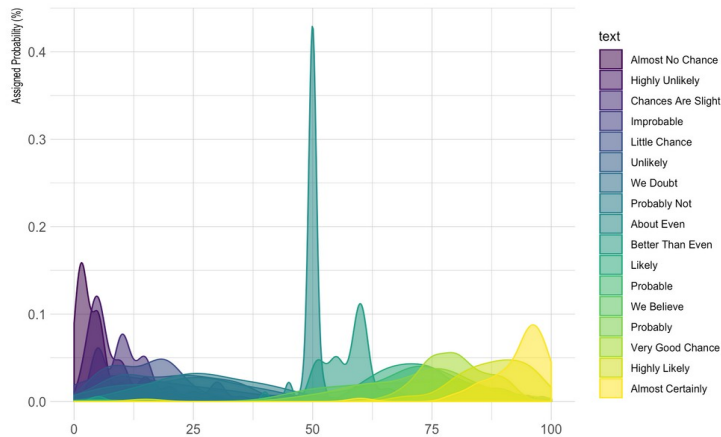


Column chart: never cut the Y-axis

Same data, and much more readable, isn't it?

VISUAL OVERLOAD

Too much information decreases readability



When it's possible: add legend directly on the chart

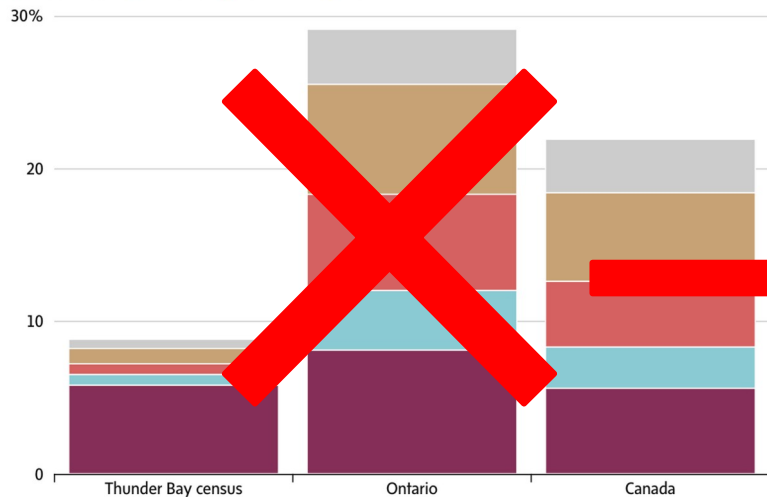
Aggregate data when it's possible

WRONG DESIGN

By convention, the time (years, months etc...) is rather horizontal. To show continuity over time, we will use lines instead.

Immigrants as a percentage of population in 2016, by period of immigration

● Before 1981 ● 1981-90 ● 1991-2000 ● 2001-10 ● 2011-16

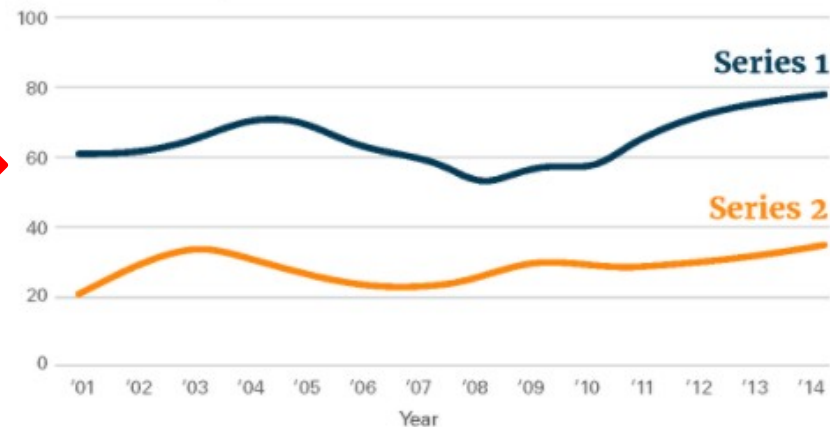


THE GLOBE AND MAIL, SOURCE: STATSCAN

DATA SHARE

Chart Title Here

(Y axis label here)

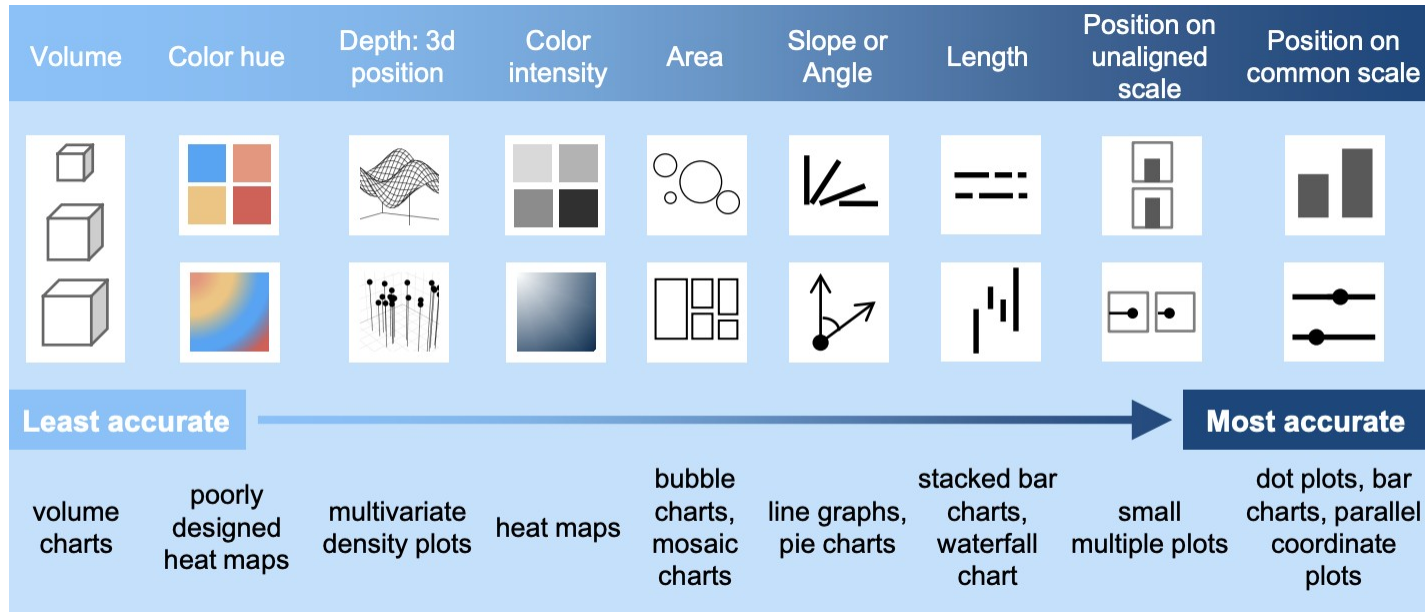


PERCEPTION OF SHAPES

Graphs are not all equal in our perception.

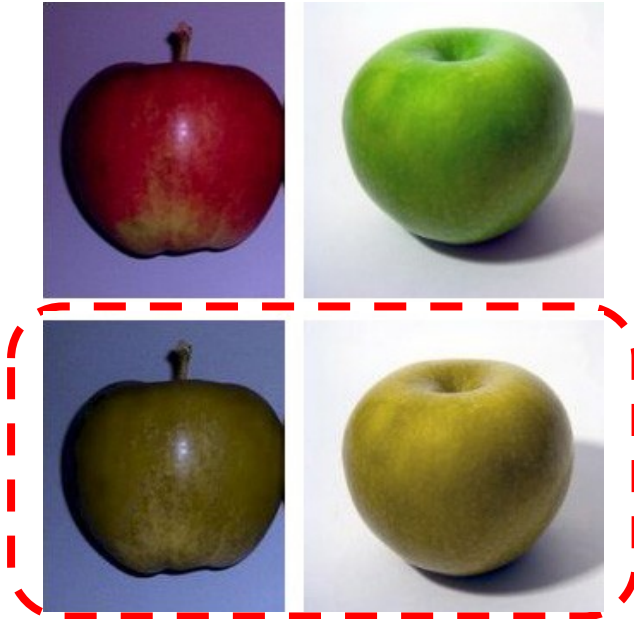
In the mid-1980s, statisticians William Cleveland and Robert McGill conducted experiments with volunteers, measuring the accuracy with which they were able to perceive quantitative information encoded by different methods (visual cues, variations in size, shape or color, etc.).

Here are their conclusions:



WRONG DESIGN

The proportion of colorblind people is about 8% in men and 0.45% in women.



What a colorblind person sees

CHARTS & ETHICS

And most importantly: always remain ethical! A dataviz is an extremely powerful tool to transmit information in an intuitive way. But you are not there to deceive or to bluff the reader, or to make the data lie. You are here to make a dataviz that will bring relevant information.

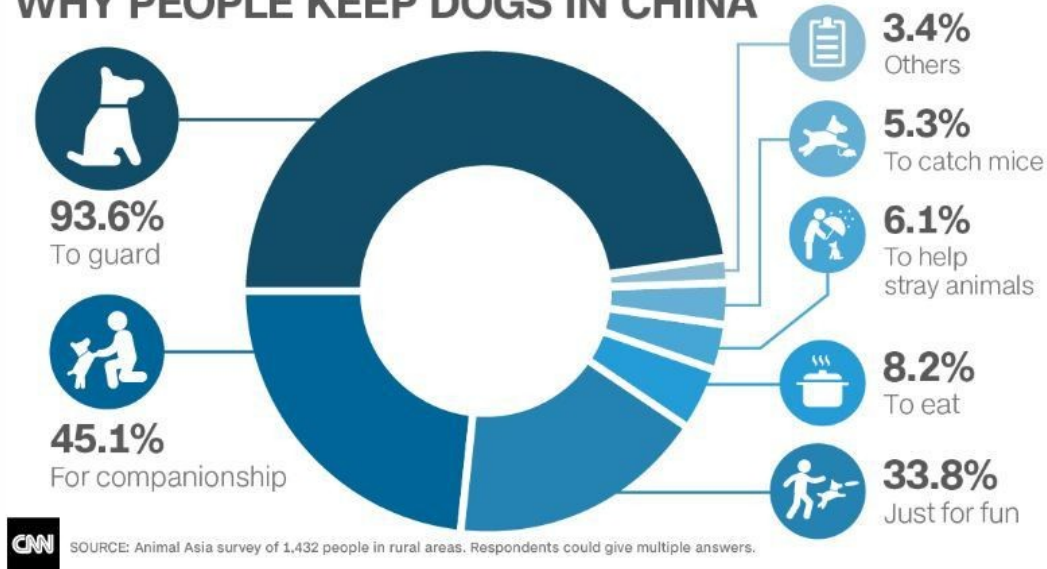
Let's go to see real bad dataviz!

Really bad examples

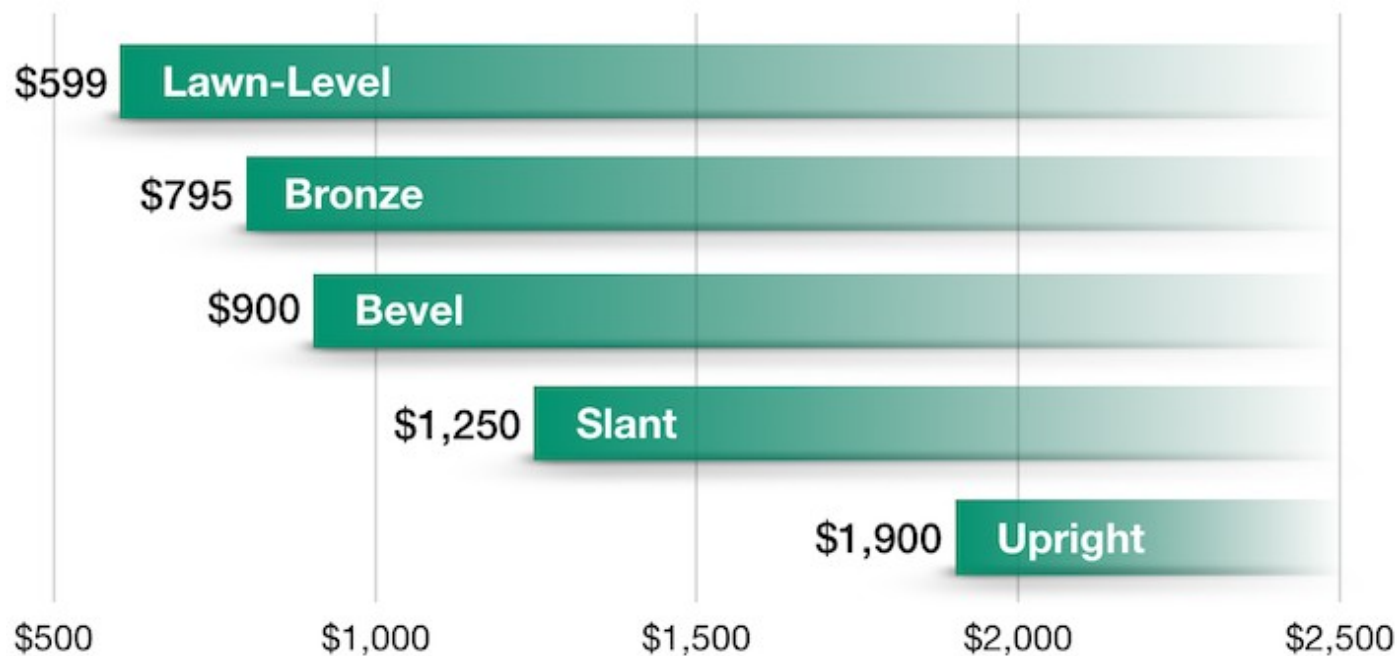
WRONG INPUT DATA

195.5% of dog lovers will find this concerning

WHY PEOPLE KEEP DOGS IN CHINA



WHAT'S WRONG HERE?



<https://www.milanomonuments.com/blog/how-much-do-memorials-headstones-and-monuments-cost-and-why>

WHAT'S WRONG HERE?

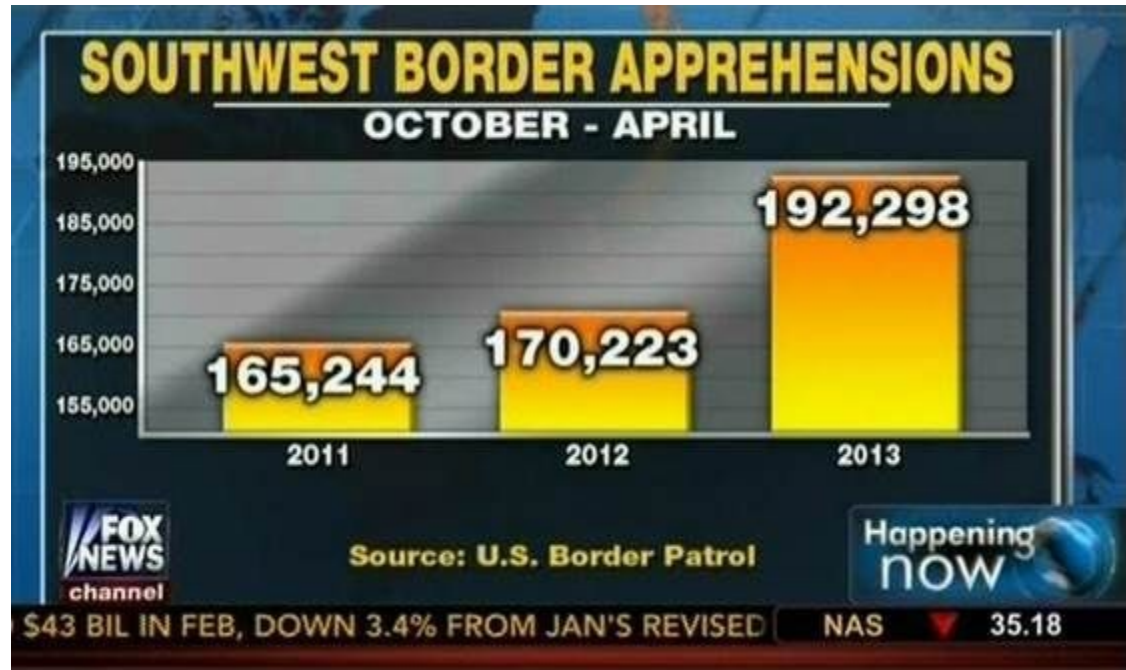


WHAT'S WRONG HERE?

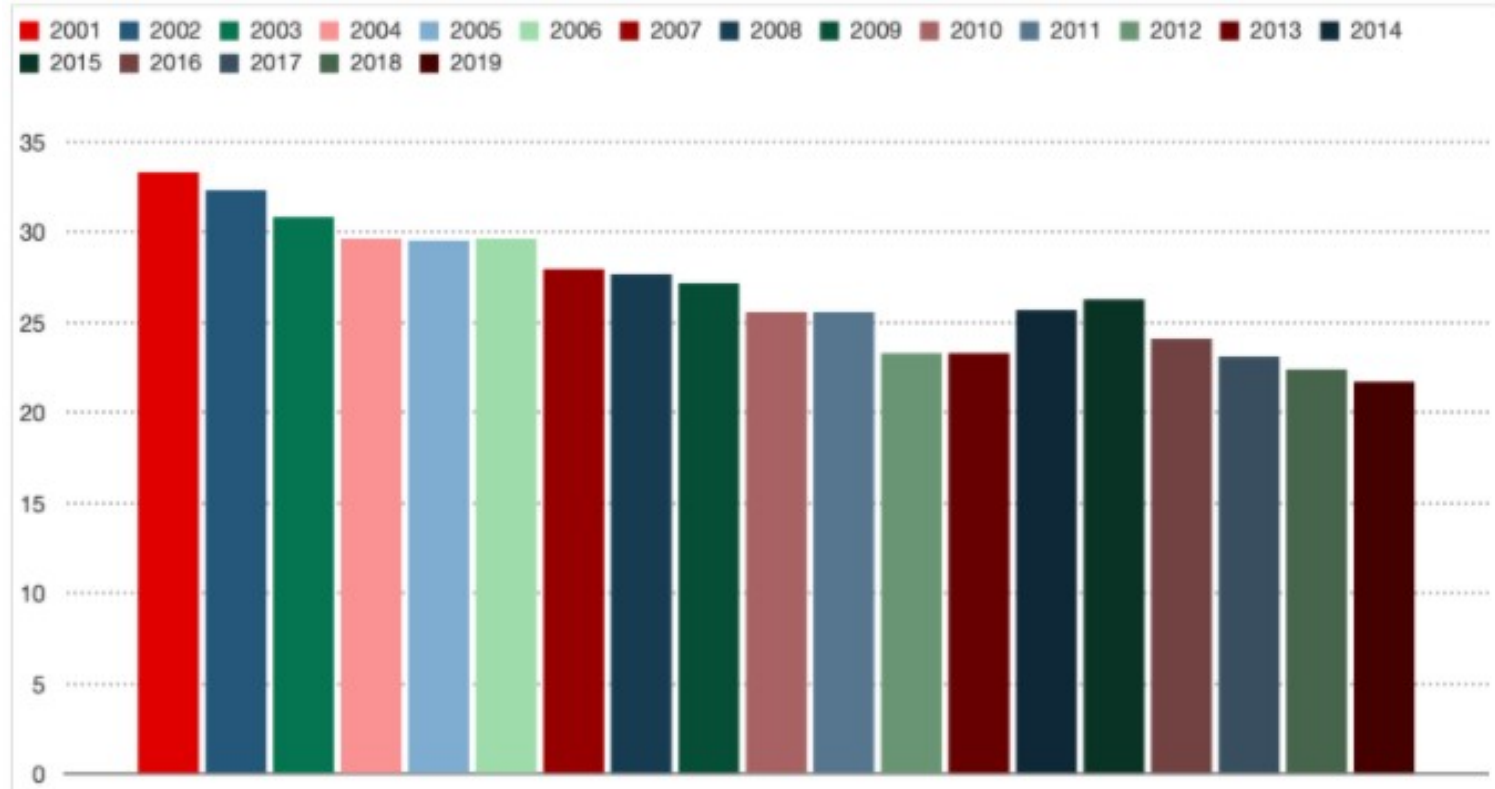


WHAT'S WRONG HERE?

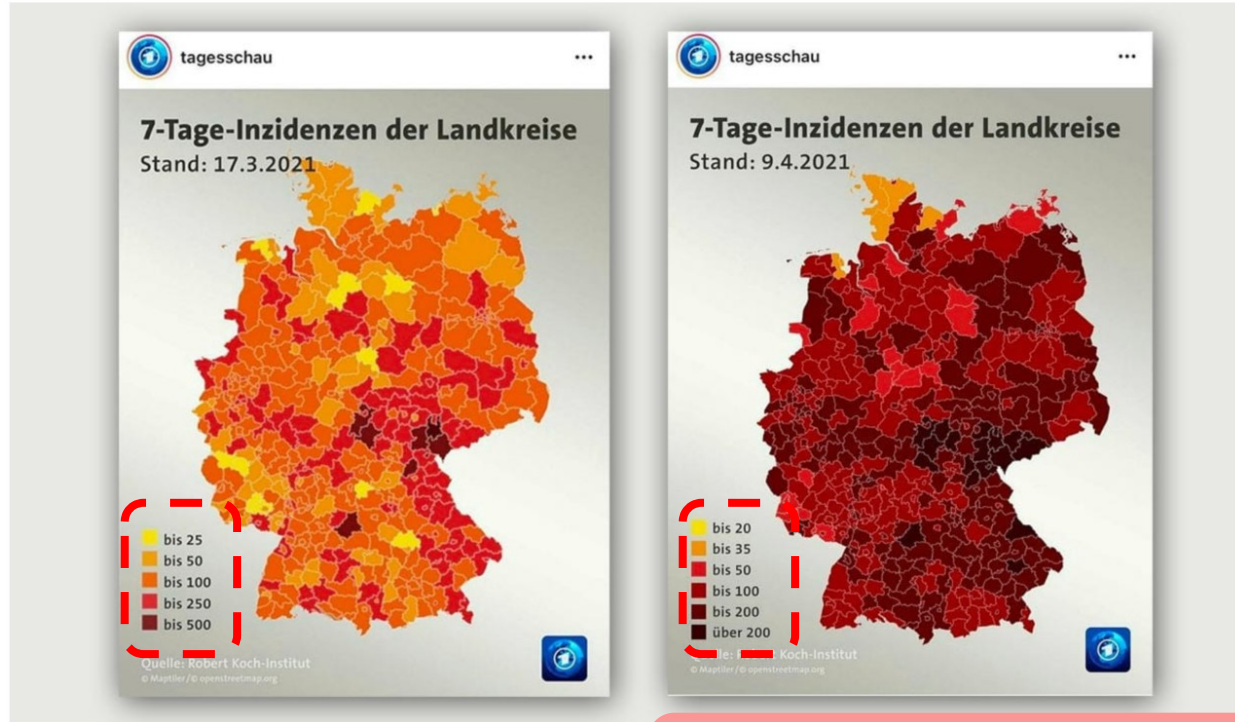
Column chart: never cut Y-axis



WHAT'S WRONG HERE?



WHAT'S WRONG HERE?



Die Tagesschau hat ihre Farbskala für Corona-Inziden

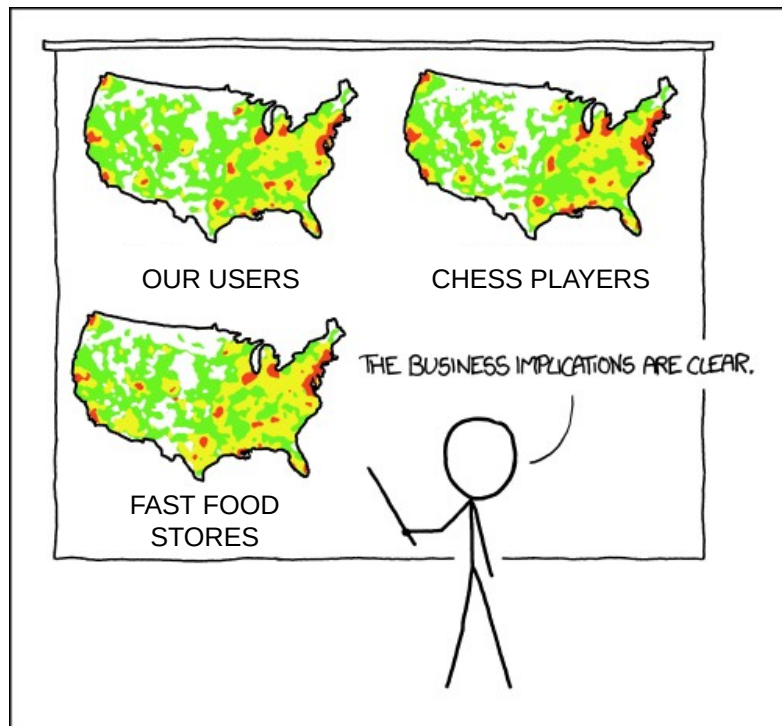
Changed the categories of the colours - so all is dark red

THE IMPORTANCE OF DATA VISUALISATION

More examples of do's and don'ts

<https://www.data-to-viz.com/caveats.html>

<https://viz.wtf>



PET PEEVE #208:
GEOGRAPHIC PROFILE MAPS WHICH ARE
BASICALLY JUST POPULATION MAPS

Storytelling

STORYTELLING

Why storytelling?

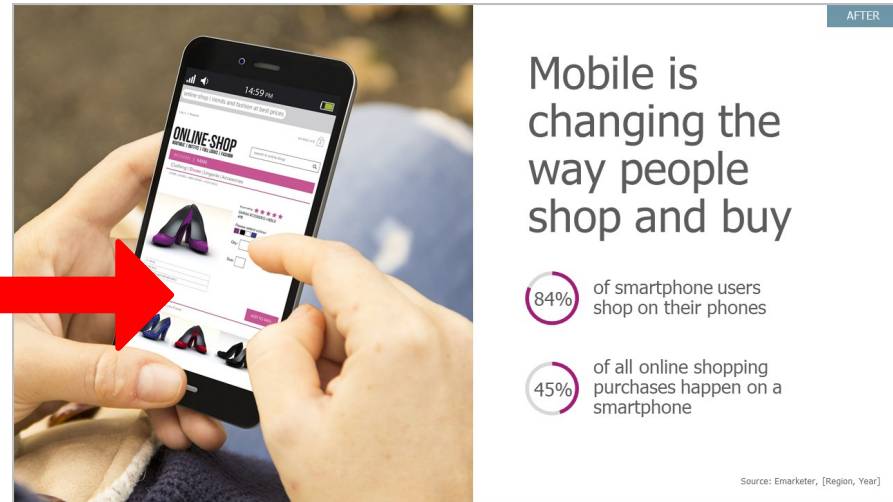
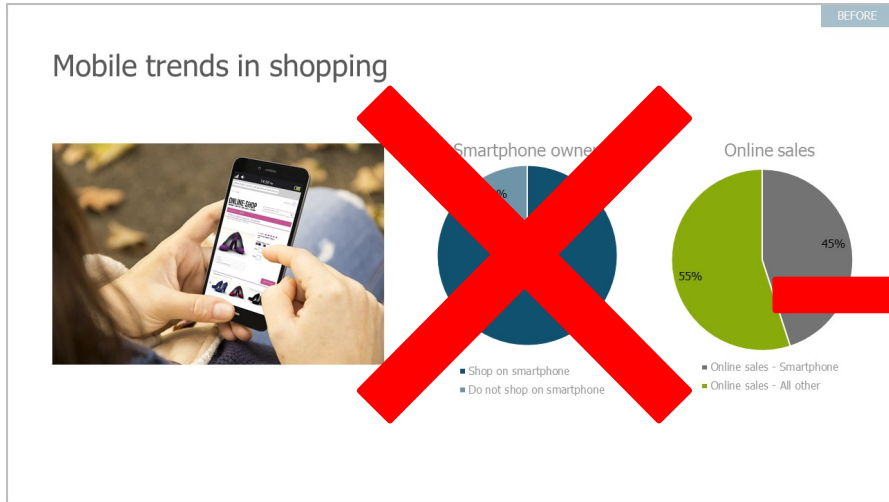
- What do I **want my audience to know** with the data I am presenting?
- How will I structure a narrative that leads **to desired action**?
- How is my data **helping drive a decision**?

How do I get started?

- What am I **trying to achieve** with the data I choose to display?
- Who is **my audience**? What do they care about? What level of data detail will they likely expect or appreciate?
- What is my **BIG Idea** -- the one thing I want my audience to know or do with my data?

STORYTELLING

Headlines with conclusion



Simple and readable content

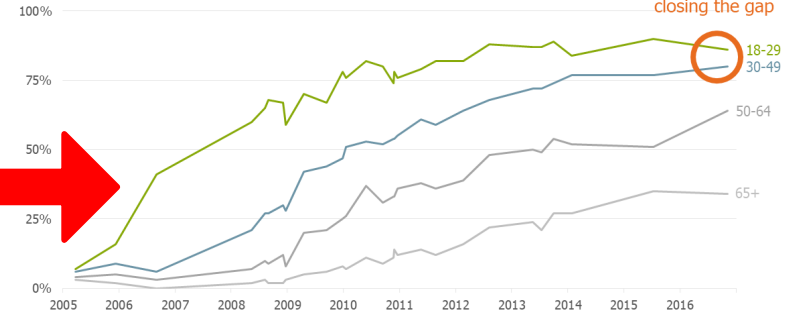
STORYTELLING

Headlines with conclusion

Social media use by age



Social media use now popular with more than just young adults



- Avoid the noise (3D effect, colors, etc...)
- Emphasis the main point
- Legend directly on the chart

Be inspired by the best!

Business Intelligence tools

Figure 1: Magic Quadrant for Analytics and Business Intelligence Platforms



Be inspired

Tableau gallery

Example:

<https://public.tableau.com/app/profile/lee0312/viz/caogao1/23>

PowerBI gallery

Example:

<https://community.powerbi.com/t5/Data-Stories-Gallery/Climate-Change-is-Real/td-p/348634>

Plotly gallery

Examples: <https://plotly.com/python/mapbox-density-heatmaps/>
<https://plotly.com/python/bubble-maps/#bubble-map-with-animation>

Streamlit gallery

Example:

<https://share.streamlit.io/streamlit/demo-uber-nyc-pickups/>