

# Constructive Criticism

---

**DATA 201: Thinking With Data**  
**Winter 2022**

Jonathan Hudson, Ph.D  
Instructor  
Department of Computer Science  
University of Calgary

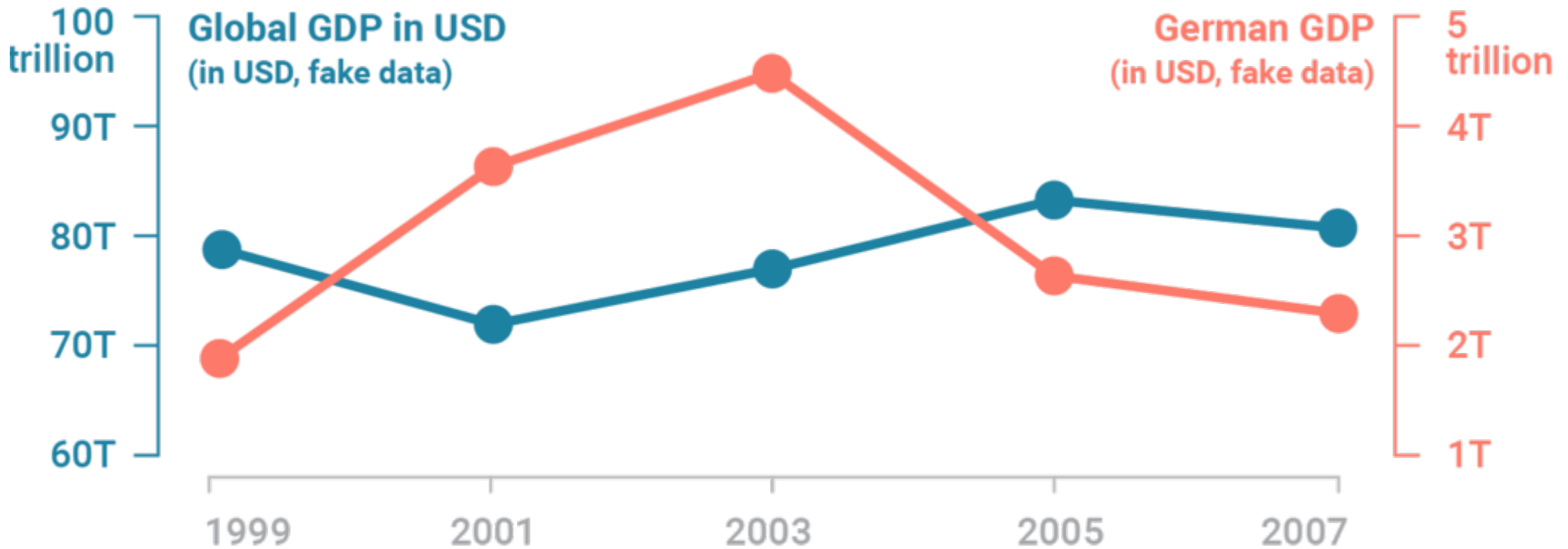
**Tuesday, March 22, 2022**



# Two Axis Charts?

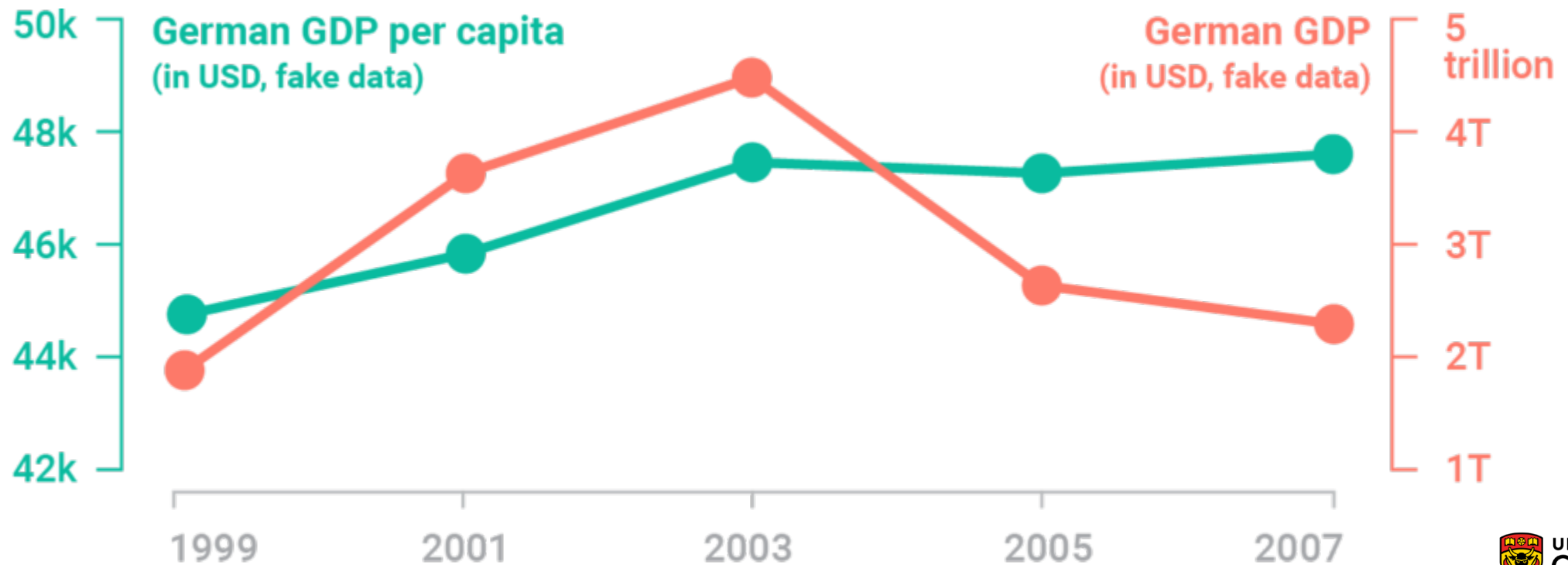
# Two Axis Charts

- Why do people use them?
- 1. Two series different magnitudes (concerning)



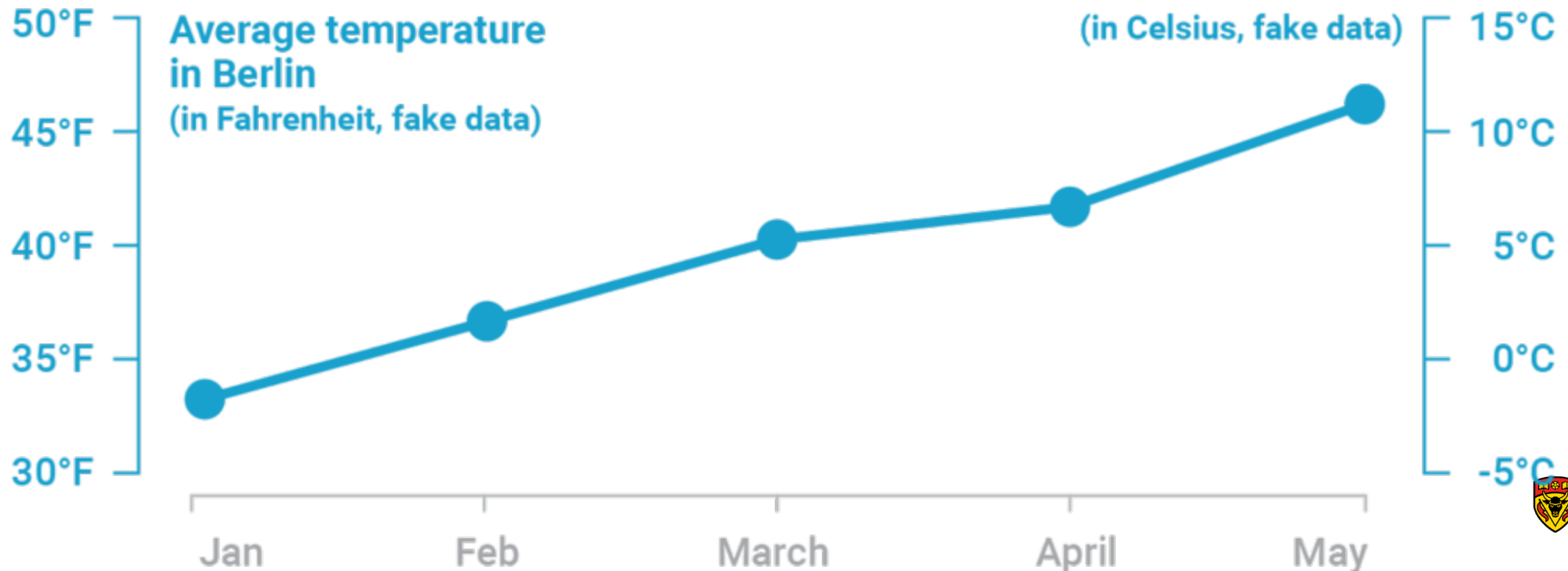
# Two Axis Charts

- Why do people use them?
  1. Two series different magnitudes
  2. Two series, relative versus absolute (benign)



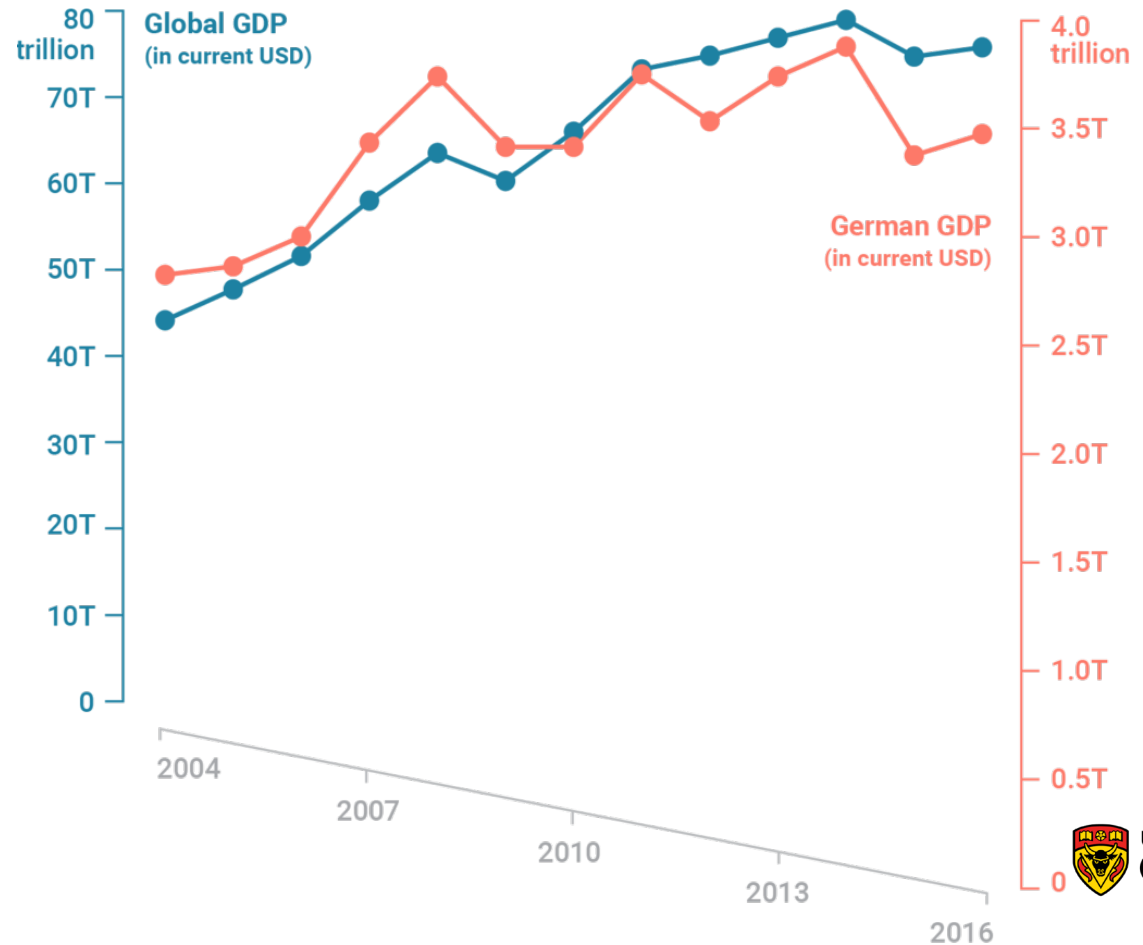
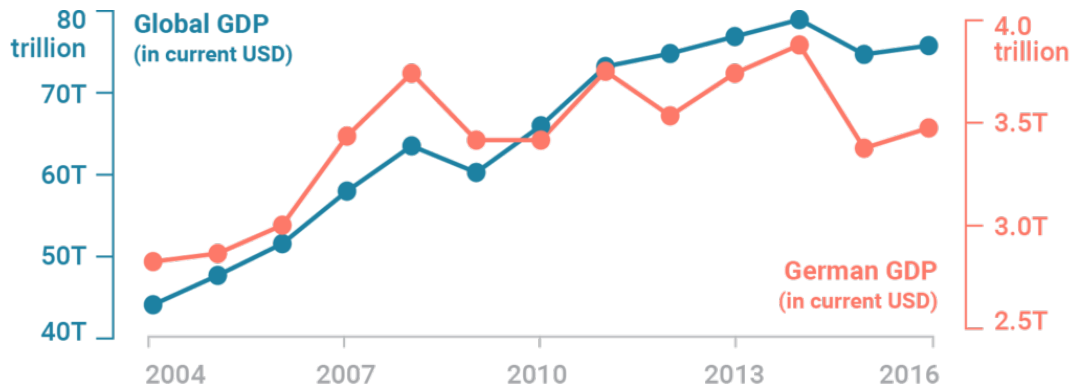
# Two Axis Charts

- Why do people use them?
  1. Two series different magnitudes
  2. Two series, relative versus absolute
  3. One series, different scales (benign)



# The problem?

- The scales of dual axis charts are arbitrary and can therefore (deliberately) mislead readers about the relationship between the two data series.



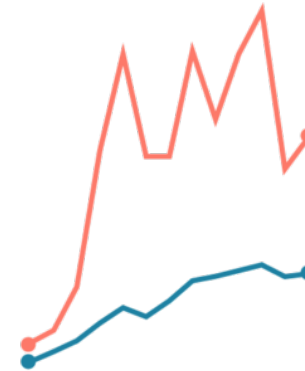
# Picks scale to change story



Orange steady,  
Blue massively increasing.



Blue steady,  
Orange increasing.



Both started at the same  
level, but Orange increased  
far more than Blue.



Both started at the same  
level, but Blue increased far  
more than Orange.

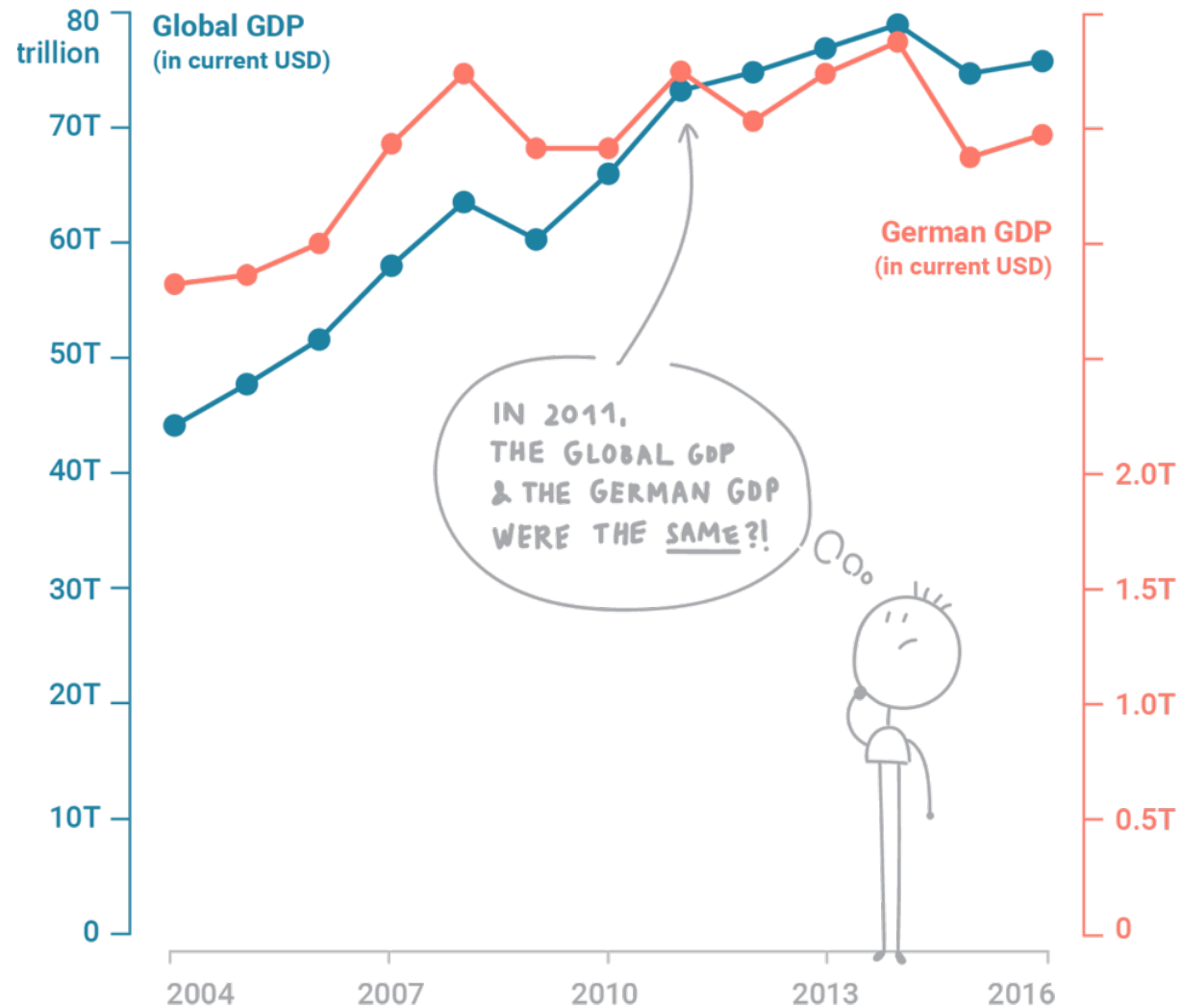


Both started with the  
same increase, then Blue  
raced to the top.



Both steady.

# Even 0 axis can cause problems

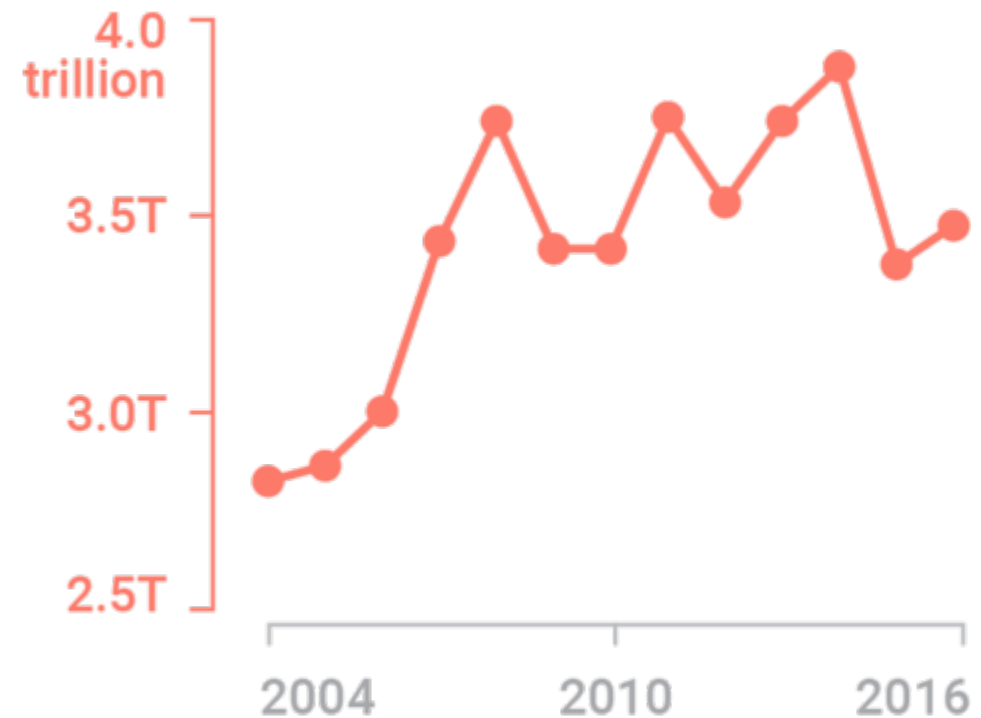
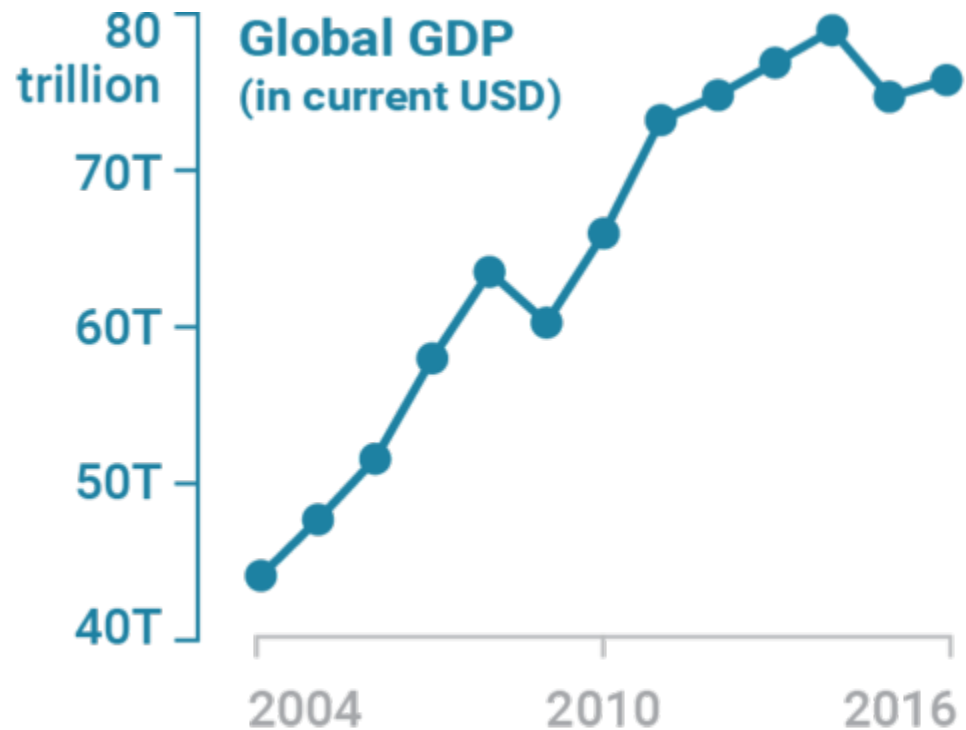




# Solutions?

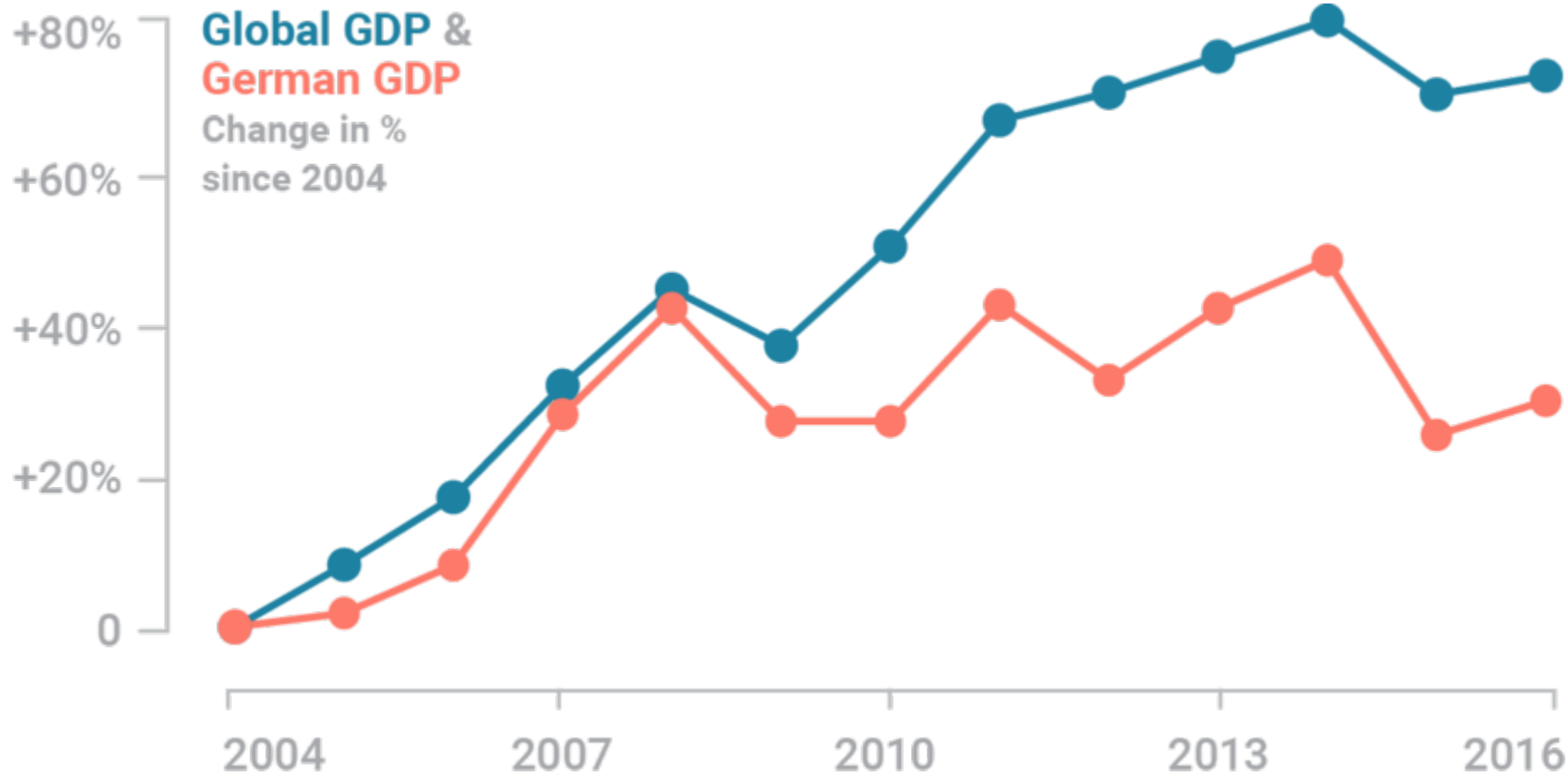
---

- Side by Side (highly recommend) Languages like R make this easy



# Solutions?

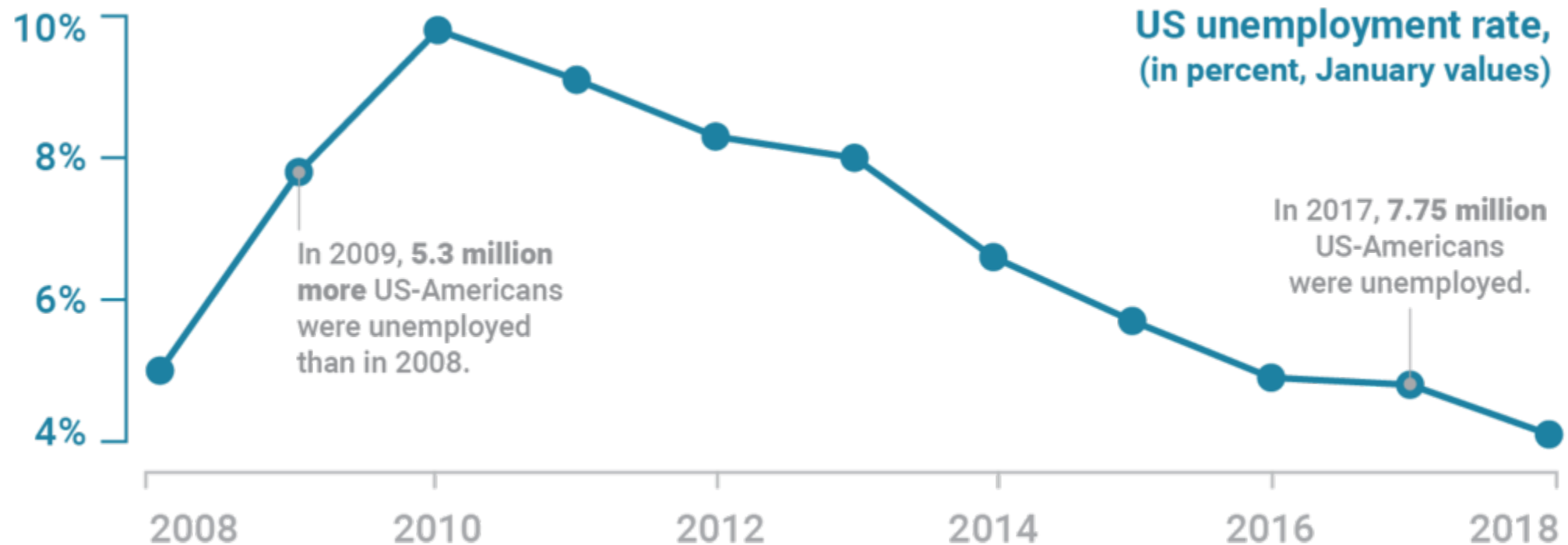
- No longer absolute, both relative (we however have lost absolute idea)



# Solutions?

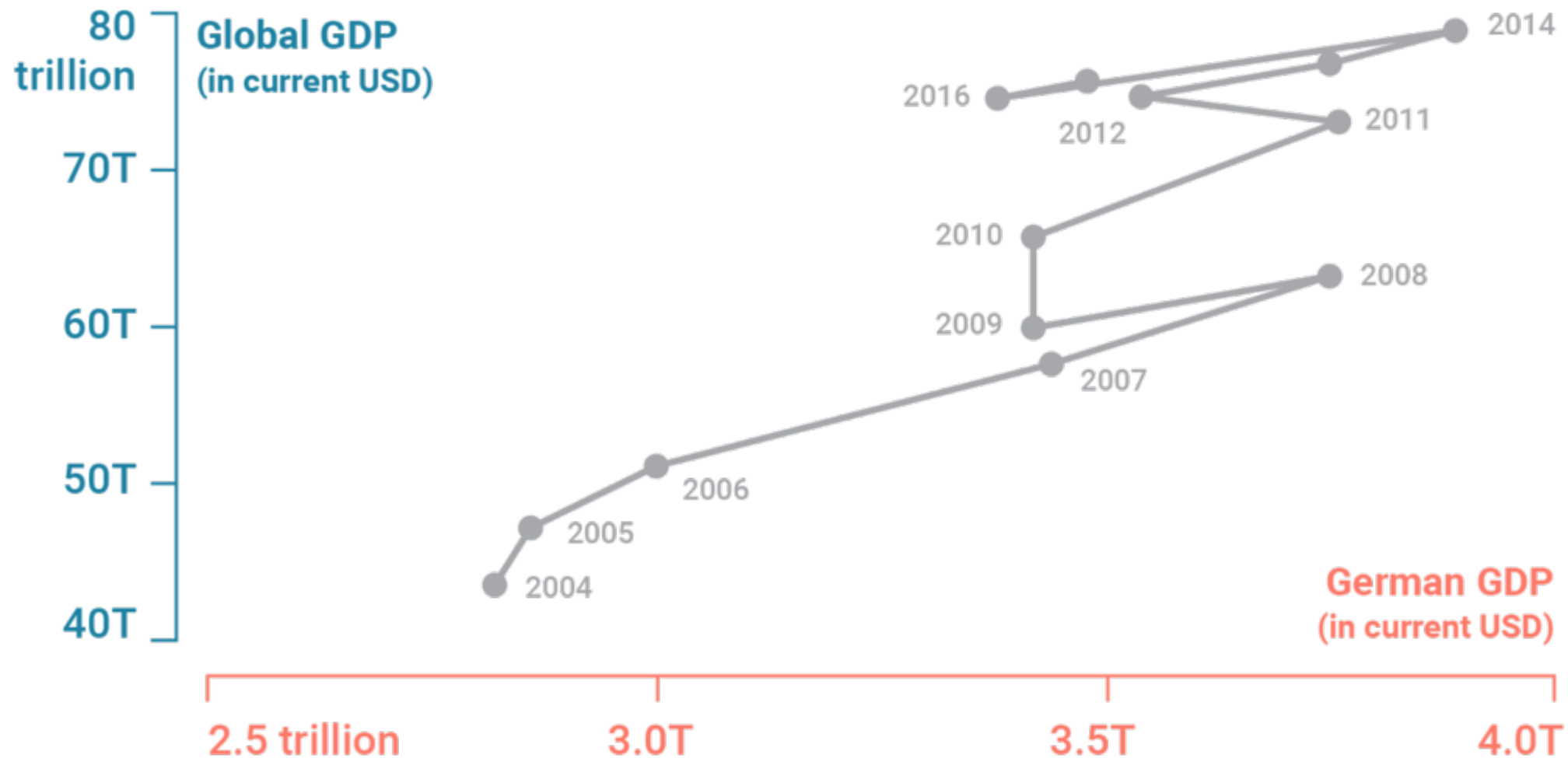
---

- Labelling in interesting data (rarely what we want, but great when it is)



# Solutions?

- Connected scatter? Not overly intuitive, but once and a while works.



# Simple Charts

# Simple Chart

- We don't mean minimal visual, or minimal data, but simple clear story (communication)
- Not animated

Non simple examples (a lot of infographics fall into here as well)

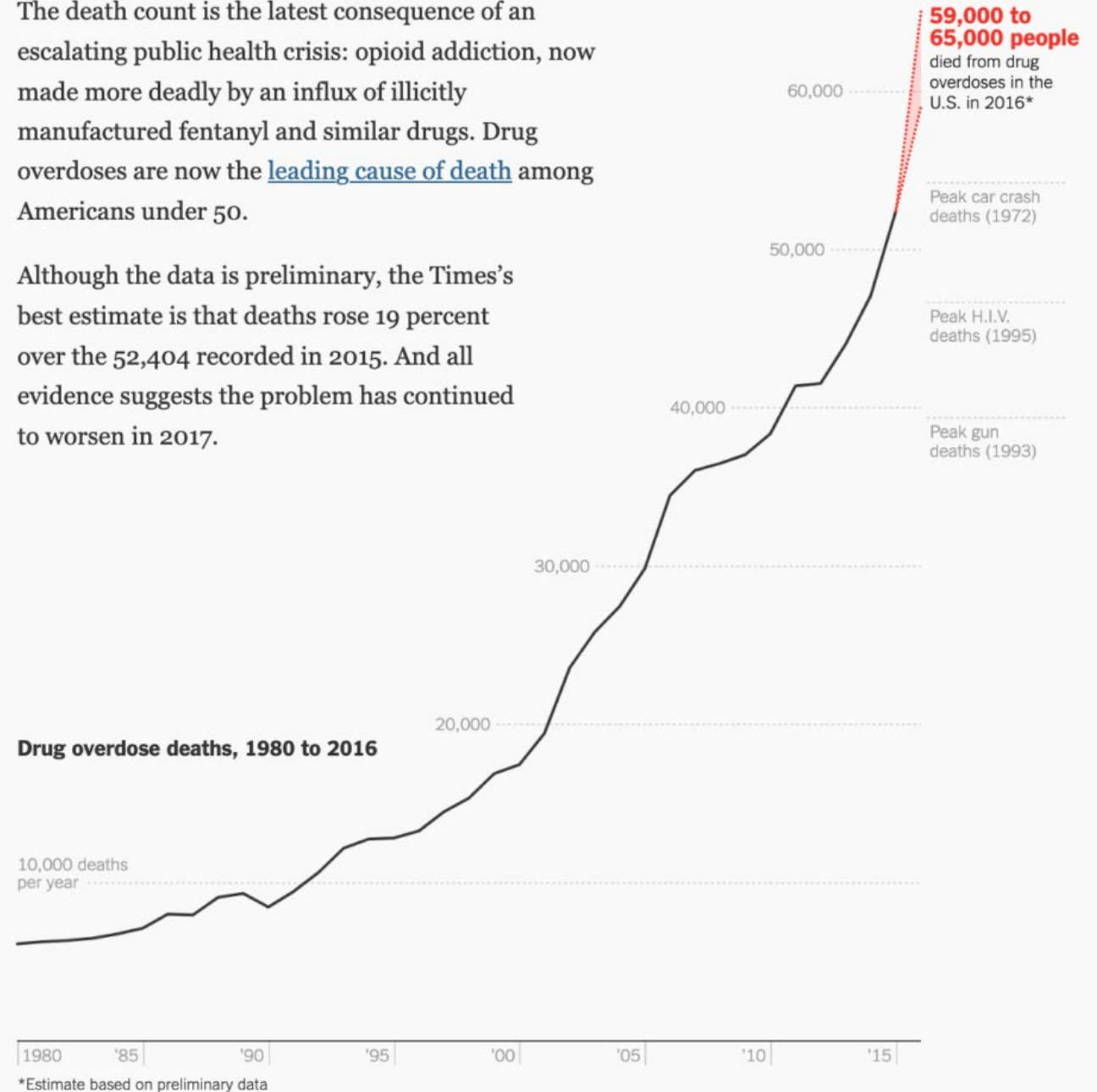


# Simple Chart

- New York Times chart of drug overdose deaths between 1980 and 2016.

The death count is the latest consequence of an escalating public health crisis: opioid addiction, now made more deadly by an influx of illicitly manufactured fentanyl and similar drugs. Drug overdoses are now the leading cause of death among Americans under 50.

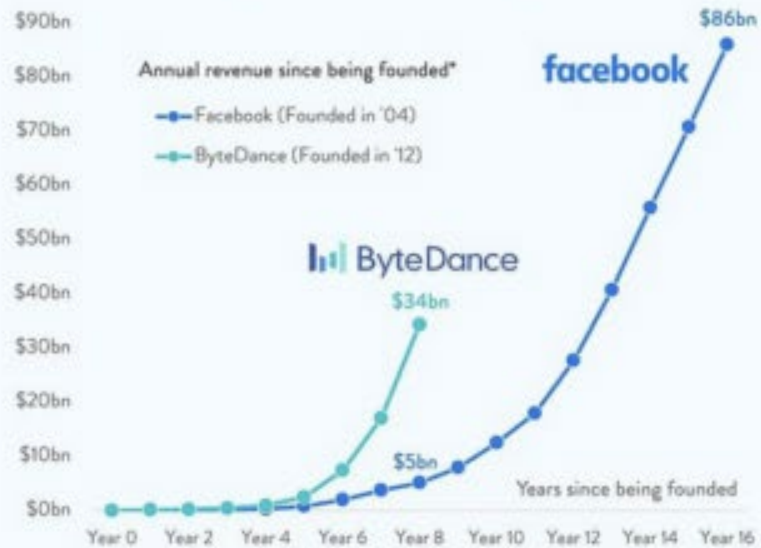
Although the data is preliminary, the Times's best estimate is that deaths rose 19 percent over the 52,404 recorded in 2015. And all evidence suggests the problem has continued to worsen in 2017.



# Simple Chart

- Chartr
- Message of each is clear, annotated, often coloured, immediate

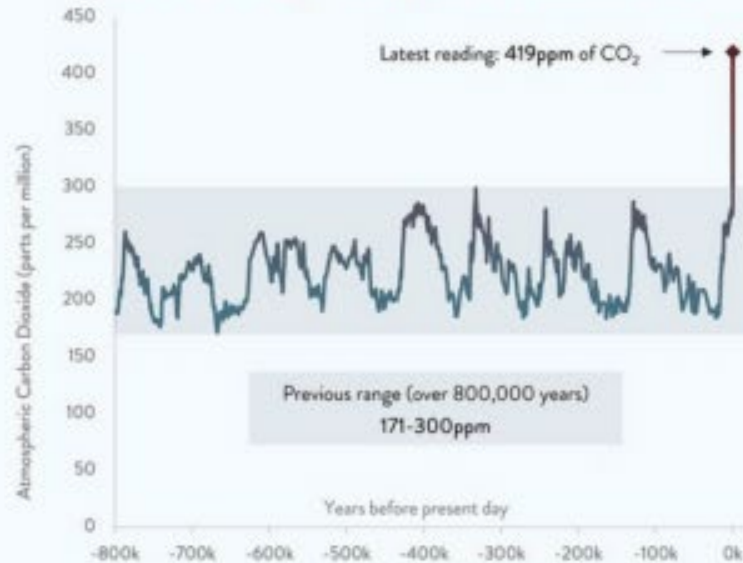
## TikTok's Parent Company ByteDance Is Growing Unbelievably Quickly



Source: WSJ, Facebook SEC Filings  
\*Years 1-3 estimated for Facebook, Years 1-5 estimated for ByteDance

chart

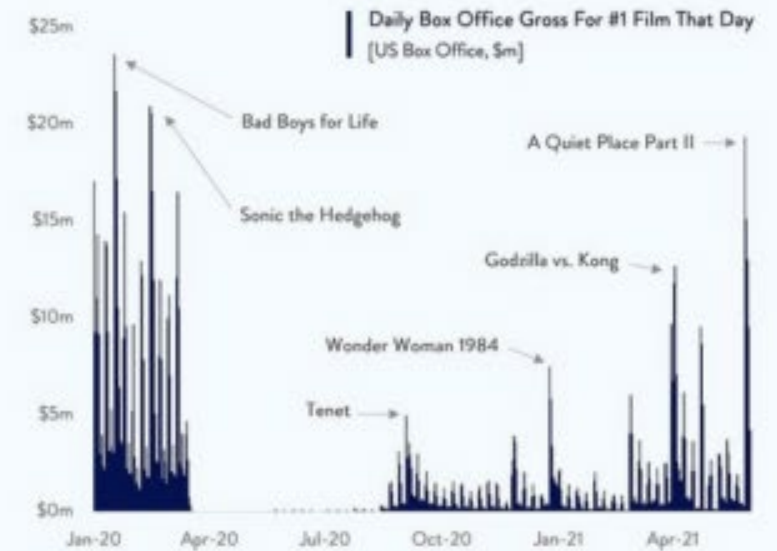
## Current Carbon Dioxide Levels Are Truly Unprecedented



Source: NOAA, Climate.gov and NCEI

chart

## The Box Office Is Back: A Quiet Place 2 Grabs \$19m On First Day



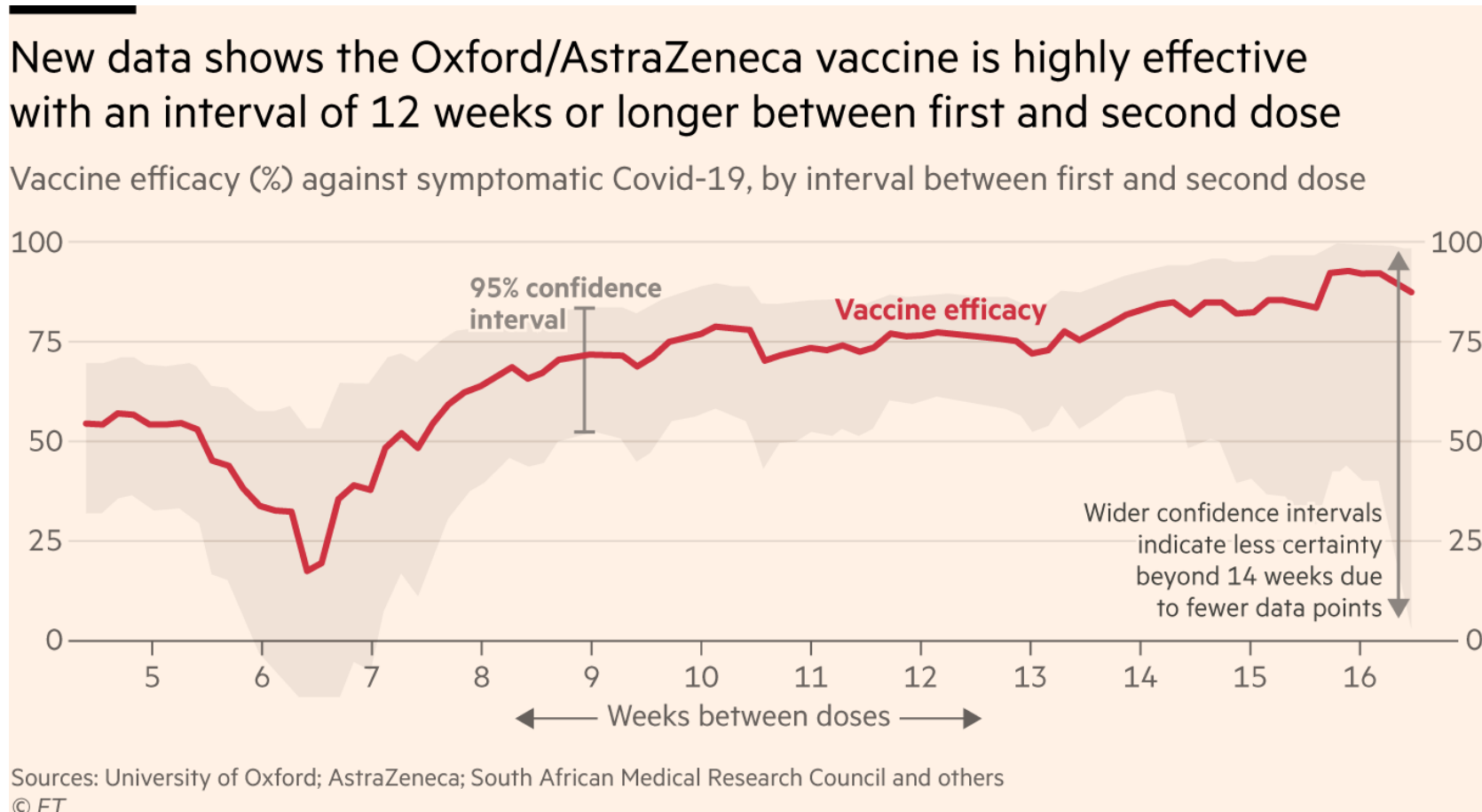
Source: Box Office Mojo

chart



# Simple Chart

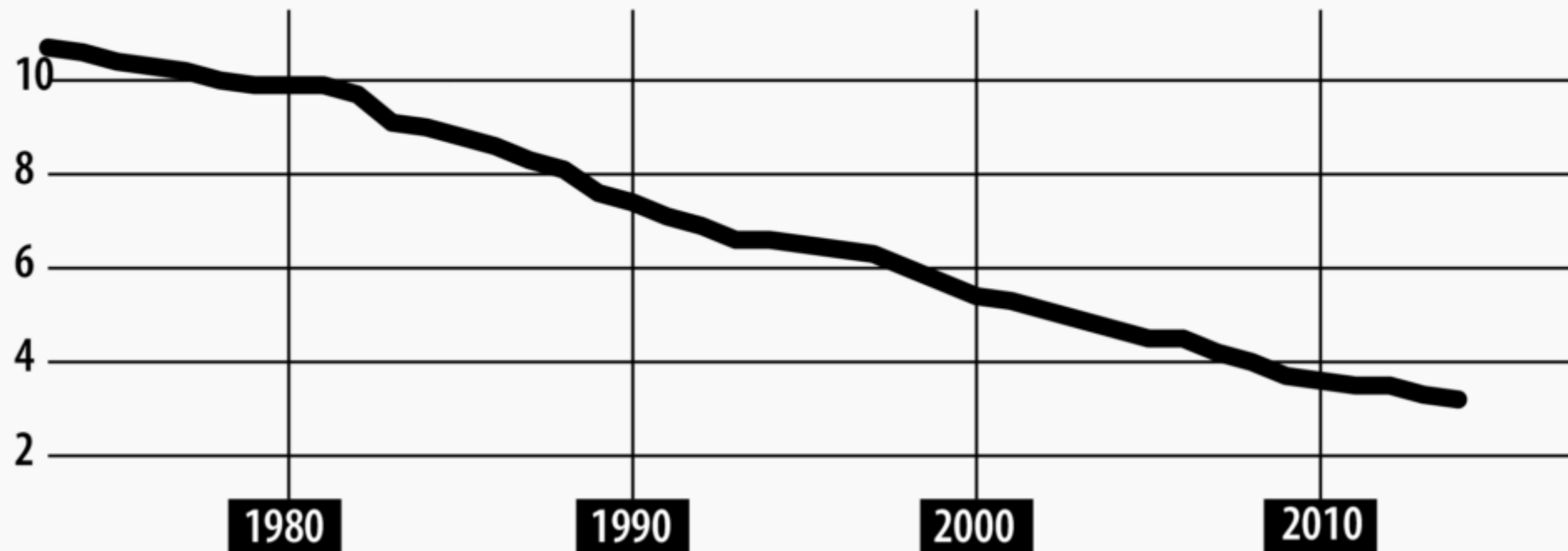
<https://twitter.com/jburnmurdoch/status/1356689724660981767/photo/1>



# Boring Simple Chart

## Sales of cigarettes per adult per day in the U.S.

National statistics, via Our World in Data



# Interesting Simple Chart

## The rise and fall of cigarette consumption in developed countries

Edit this chart

Sales of cigarettes per adult per day, in selected countries. Figures include manufactured cigarettes, as well as an estimated number of hand-rolled cigarettes, per adult (ages 15+) per day.

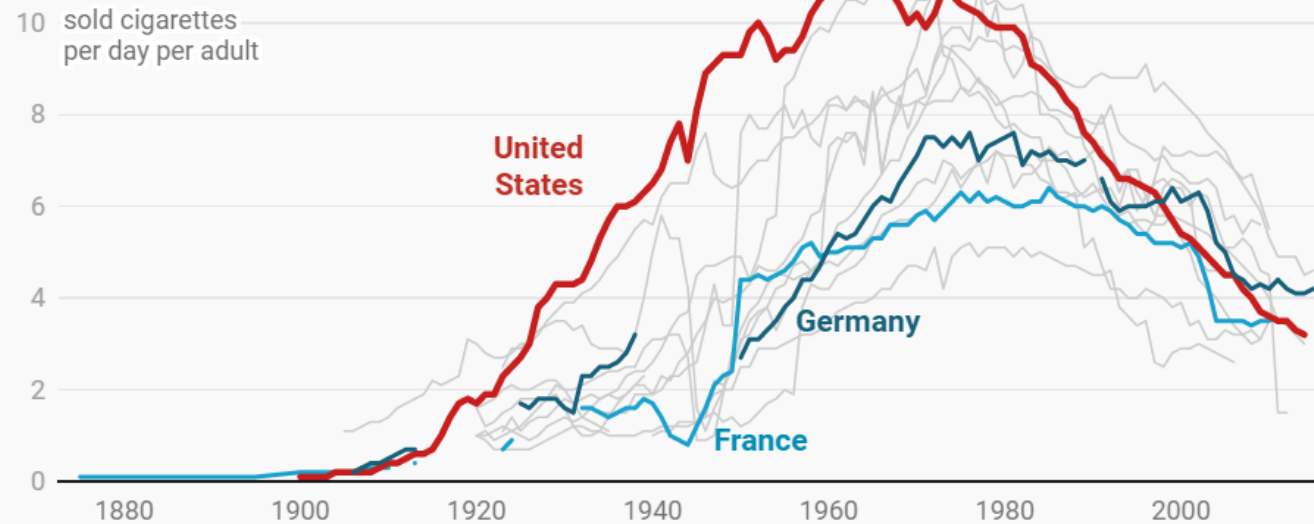


Chart: Lisa Charlotte Rost • Source: [National statistics](#), via [Our World in Data](#) • [Get the data](#)

# Simple Chart

- Other lessons (mix-up chart types)
- Diversity continues interest and can help stretch the readers understanding to consider information from different angles
- Clarity over complexity

## Three things people spent money on in 2020



Hi, I'm Jakob, a software engineer at Detoxifier. In this week's article, I'll talk about everyone's favorite thing: money. And also about everyone's least favorite thing: the year 2020.

There are many ways to tell the story of last year. Each person has their own experience. How do we describe the whole as well as its parts? Today, I will try by asking one particular question: What were people spending their money on in 2020? I will give three different answers, to tell three different stories.

### Computers

One thing people apparently liked to (or had to?) spend their money on last year were PCs. While in 2019, manufacturers shipped 263 million personal computers worldwide, in 2020, the number increased by 10% and even more at the beginning of 2021. Whereas in the years before, the changes were rather modest (3% in 2017, 1% in 2018, +1% in 2019). The current jump in computer sales is a story of digital transformation – work from home, remote learning, gaming.

#### People are buying more PCs

Number of personal computers shipped from manufacturers



Source: Statista. [Get the data](#). Created with Detoxifier.

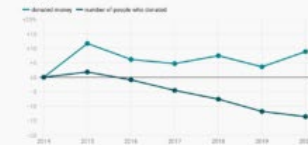
### Charity

Another area that saw an increase in spending was charity donations. In Germany, people donated 5.4 billion Euros last year, which is 5.1% more than the year before. This increase is not as big as the one due to the refugee crisis of 2015, when donations rose by 11% compared to 2014. The money spent on charity tells a story of altruism. In Germany, people donated more to humanitarian causes, refugees, children and youth organizations, animal shelters, and cultural heritage organizations, less to sports.

It's also interesting to notice that although the amount of money donated increased, the number of people who donated decreased. This has been a trend for several years – 33% of people donated in 2014, but only 28.5% in 2020. We can see that as a story of inequality – some are more able to give than others.

#### In Germany, fewer people donate, but more money gets donated

Percentage change of donated money and donating people between 2014 and 2020, with 2014 as a base



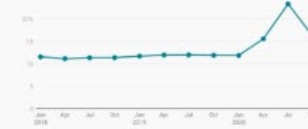
Source: Statista. [Get the data](#). Created with Detoxifier.

### Nothing

It wouldn't be a *bit*icle if there weren't one weird item. So in this tradition, my final answer to "What were people spending their money on?" is... nothing. People didn't spend their whole income. Household savings raised sharply in 2020. This tells us a story of precaution, risk of future unemployment, postponement of big purchases, but also of not being able to go on vacation.

#### People in the EU are saving more money

Percentage of income that households save



Source: Trading Economics. [Get the data](#). Created with Detoxifier.

That's it for my Weekly Chart. Now let's sit back, relax, and see which of these money-related developments were just temporary – and which are here to stay. Next Thursday, our support engineer Aya will create a Weekly Chart. We'll see you then.

## Three things people spent money on in 2020



Hi, I'm Jakob, a software engineer at Detoxifier. In this week's article, I'll talk about everyone's favorite thing: money. And also about everyone's least favorite thing: the year 2020.

There are many ways to tell the story of last year. Each person has their own experience. How do we describe the whole as well as its parts? Today, I will try by asking one particular question: What were people spending their money on in 2020? I will give three different answers, to tell three different stories.

### Computers

One thing people apparently liked to (or had to?) spend their money on last year were PCs. While in 2019, manufacturers shipped 263 million personal computers worldwide, in 2020, the number increased by 10% and even more at the beginning of 2021. Whereas in the years before, the changes were rather modest (3% in 2017, 1% in 2018, +1% in 2019). The current jump in computer sales is a story of digital transformation – work from home, remote learning, gaming.

#### People are buying more PCs

Number of personal computers shipped from manufacturers



Source: Statista. [Get the data](#). Created with Detoxifier.

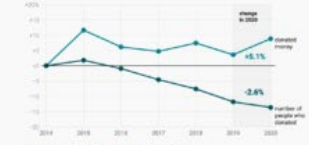
### Charity

Another area that saw an increase in spending was charity donations. In Germany, people donated 5.4 billion Euros last year, which is 5.1% more than the year before. This increase is not as big as the one due to the refugee crisis of 2015, when donations rose by 11% compared to 2014. The money spent on charity tells a story of altruism. In Germany, people donated more to humanitarian causes, refugees, children and youth organizations, animal shelters, and cultural heritage organizations, less to sports.

It's also interesting to notice that although the amount of money donated increased, the number of people who donated decreased. This has been a trend for several years – 33% of people donated in 2014, but only 28.5% in 2020. We can see that as a story of inequality – some are more able to give than others.

#### In Germany, fewer people donate, but more money gets donated

Percentage change of donated money and donating people between 2014 and 2020, with 2014 as a base



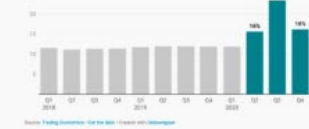
Source: Statista. [Get the data](#). Created with Detoxifier.

### Nothing

It wouldn't be a *bit*icle if there weren't one weird item. So in this tradition, my final answer to "What were people spending their money on?" is... nothing. People didn't spend their whole income. Household savings raised sharply in 2020. This tells us a story of precaution, risk of future unemployment, postponement of big purchases, but also of not being able to go on vacation.

#### People in the EU are saving more money

Percentage of income that households save



Source: Trading Economics. [Get the data](#). Created with Detoxifier.

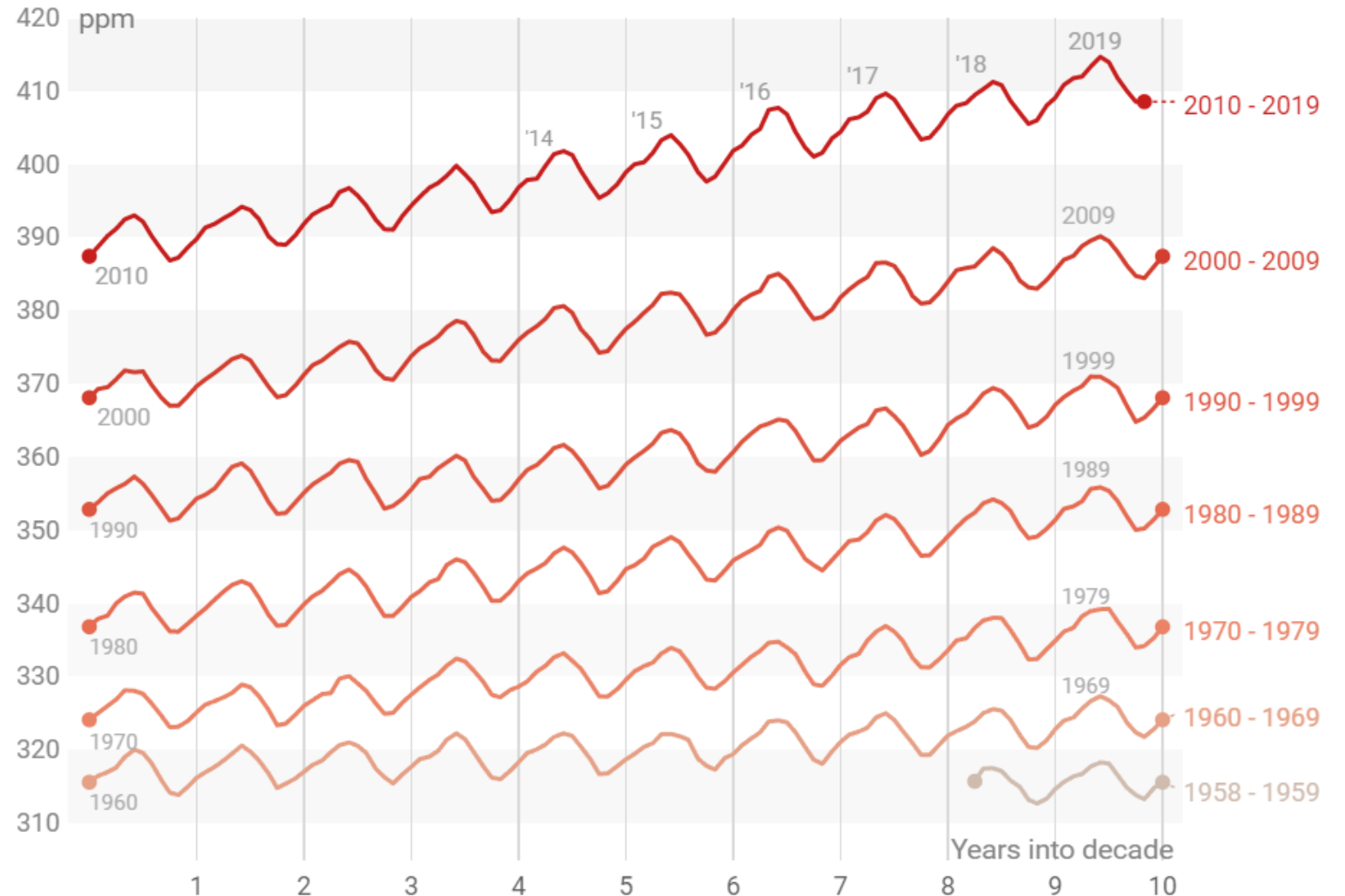
That's it for my Weekly Chart. Now let's sit back, relax, and see which of these money-related developments were just temporary – and which are here to stay. Next Thursday, our support engineer Aya will create a Weekly Chart. We'll see you then.

# Simple Chart

- Examples of clarity over complexity

## Six Decades of Carbon Dioxide Concentration in the Atmosphere

CO2 concentration in parts per million\* (ppm). Each line represents one decade, from 1958 to 2019.

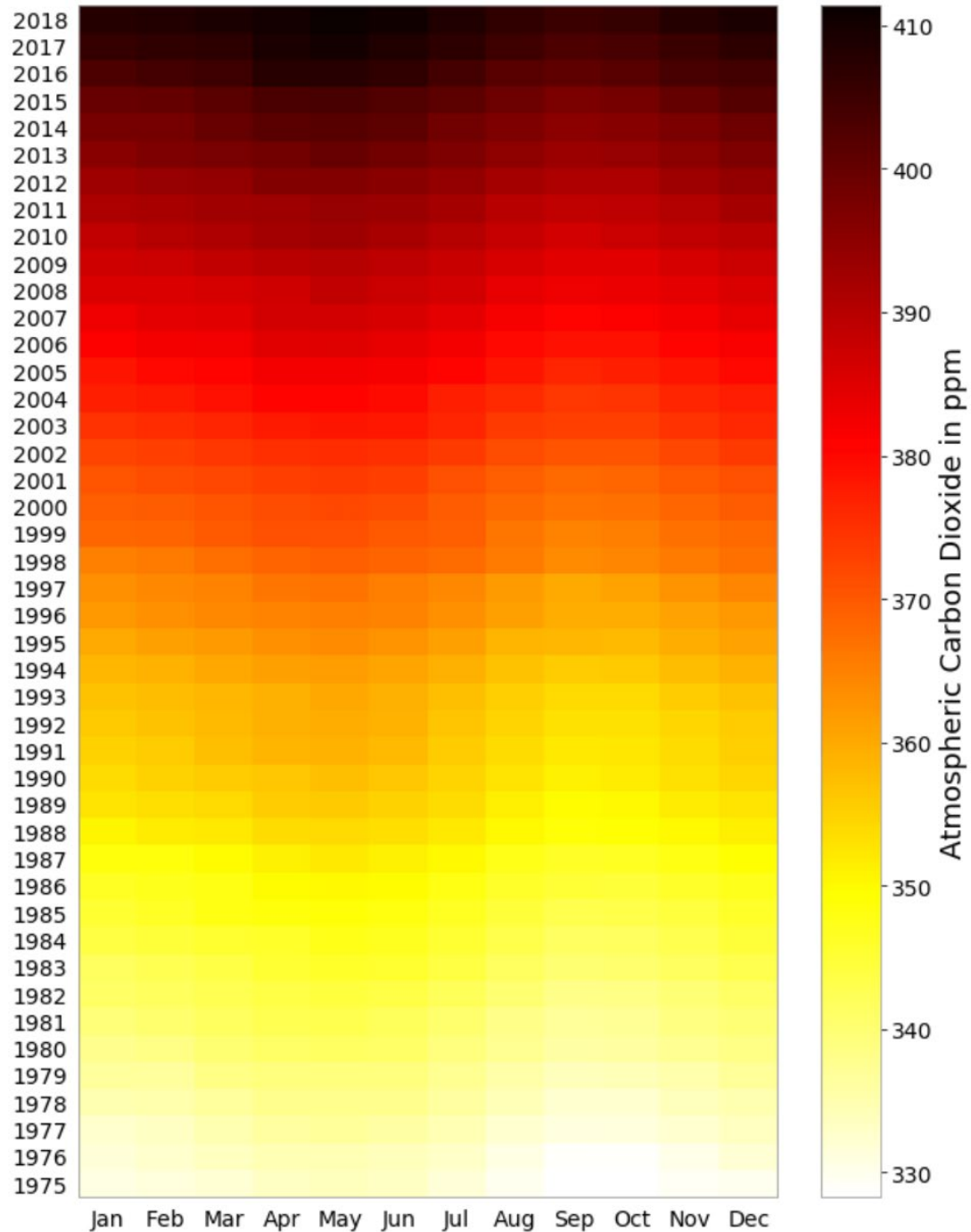


\*The mole fraction of CO2, expressed as parts per million (ppm) is the number of molecules of CO2 in every million molecules of dried air (water vapor removed).

Source: [National Oceanic & Atmospheric Adm. \(NOAA\)](#) • [Get the data](#)

# Simple Chart

- Examples of clarity over complexity

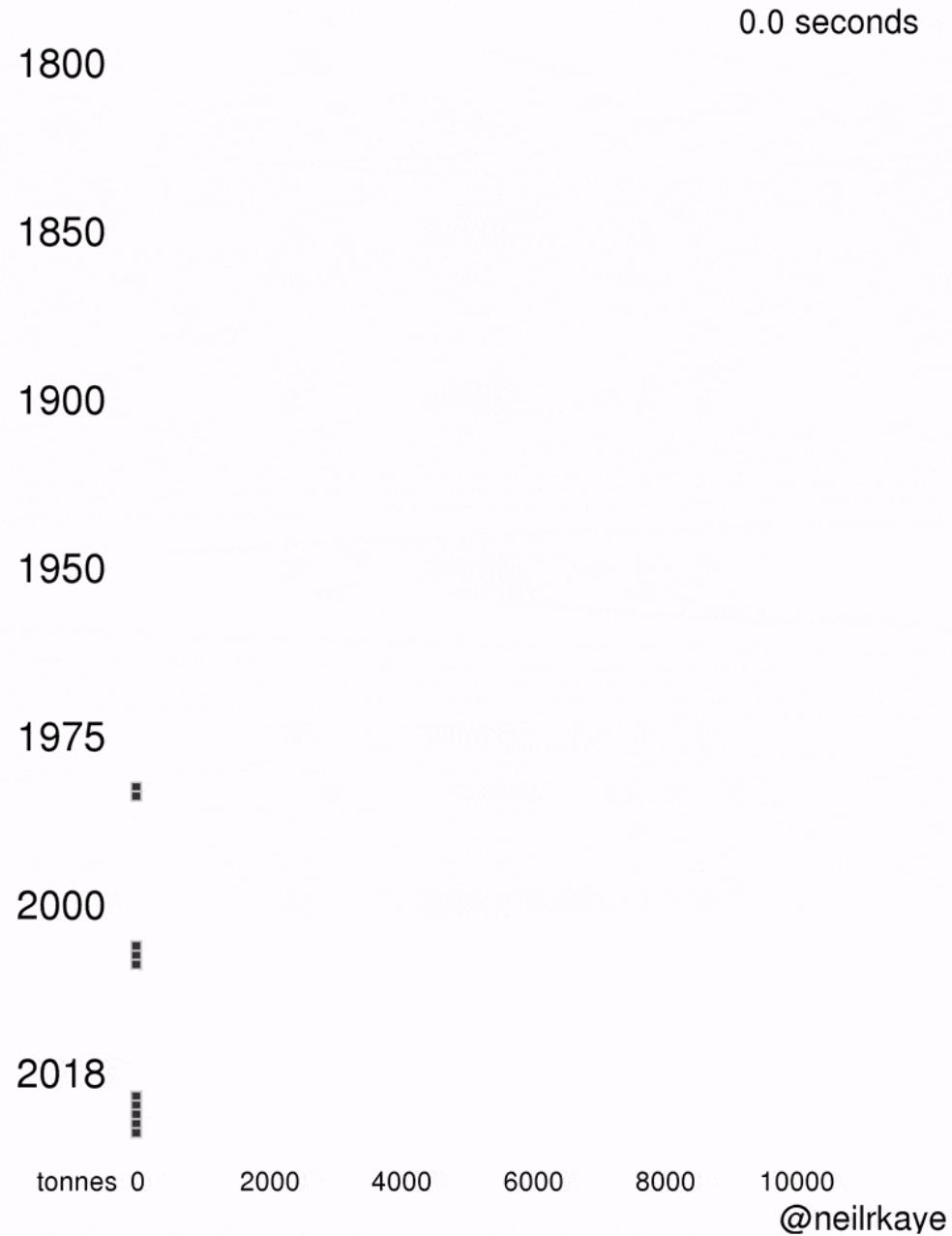


“Visualizing Atmospheric Carbon Dioxide” by Giuseppe Vettigli, 2019. [Learn how he created the visualization in Python here.](#)

# Simple Chart

- Examples of clarity over complexity

Real time speed of global fossil fuel emissions  
(each box represents 10 tonnes of CO<sub>2</sub>)

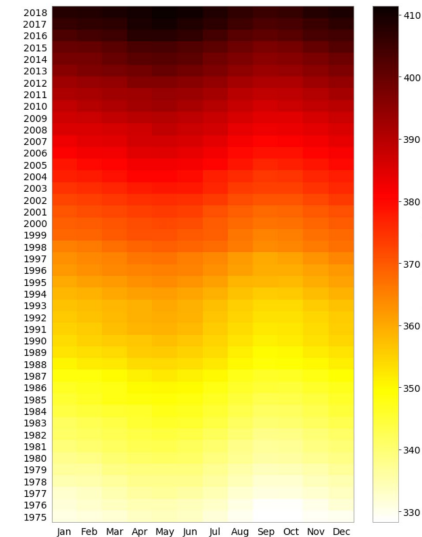
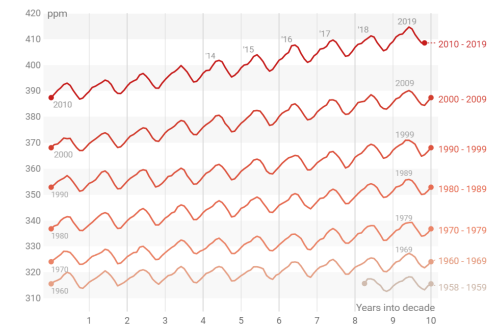


# Choice?

1. Details are best in first
  2. Simplistic in second (for your older relative?)
  3. Highlight single decade in comparison in third
- You can judge different visualizations based on different goals.
  - It's not always logical to judge them using the same criteria.
  - Ultimately it is about the audience and what you want them to understand.
  - Hiding details can be a criticism and a benefit depending on audience.

Six Decades of Carbon Dioxide Concentration in the Atmosphere

CO2 concentration in parts per million\* (ppm). Each line represents one decade, from 1958 to 2019.

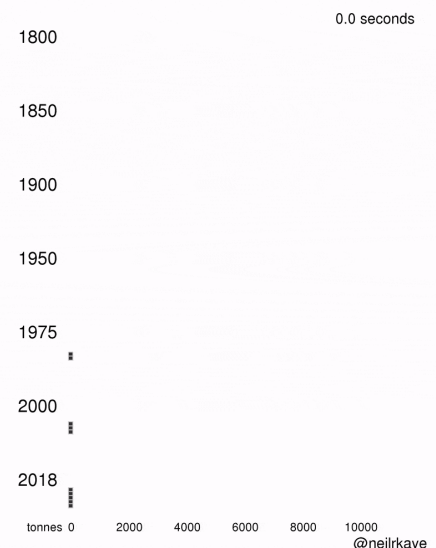


"Visualizing Atmospheric Carbon Dioxide" by Giuseppe Vettigli, 2019. Learn how he created the visualization in Python here.

\*mole fraction of CO<sub>2</sub>, expressed as parts per million (ppm) is the number of molecules of CO<sub>2</sub> in every million molecules of dried air (water vapor removed).

cc: National Oceanic & Atmospheric Adm. (NOAA) • Get the data

Real time speed of global fossil fuel emissions (each box represents 10 tonnes of CO<sub>2</sub>)





# Gendering

# Alternatives to Pink/Blue?

Rather common still

There is argument that traditional audience will respond to this scheme (due to historical gendering)

Men	Women	<b>The Pudding, 2017:</b> "Film Dialogue"
Men	Women	<b>The Pudding, 2017:</b> "She Giggles, He Gallops"
Men	Women	<b>Bloomberg, 2016:</b> "This Chart Shows Who Marries CEOs, Doctors, Chefs and Janitors"
Men	Women	<b>NYT, 2015:</b> "The Changing Nature of Middle-Class Jobs"
Men	Women	<b>NYT, 2017:</b> "The Words Men and Women Use When They Write About Love"
Men	Women	<b>Wall Street Journal, 2016:</b> "What's Your Pay Gap?"
Men	Women	<b>DailyMail, 2018</b>
Men	Women	<b>ZEIT 2016</b>
Men	Women	<b>ZEIT 2018</b>

# Alternatives to Pink/Blue?

These colours come with the whole gender stereotype baggage.

When we create a chart with pink & blue, we endorse gender stereotypes.

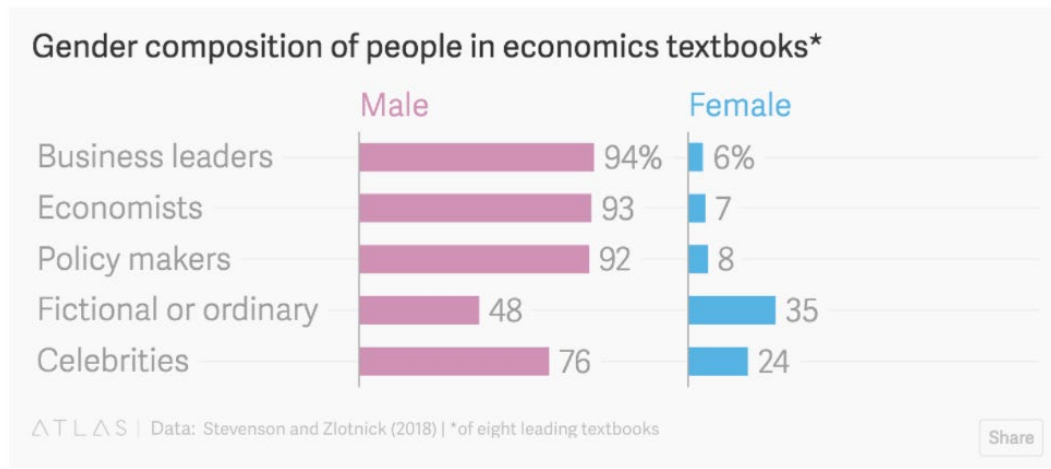
Especially when we chart gender (pay) gaps, that's often the opposite of what we want to achieve.

Men	Women	<b>The Pudding, 2017:</b> "Film Dialogue"
Men	Women	<b>The Pudding, 2017:</b> "She Giggles, He Gallops"
Men	Women	<b>Bloomberg, 2016:</b> "This Chart Shows Who Marries CEOs, Doctors, Chefs and Janitors"
Men	Women	<b>NYT, 2015:</b> "The Changing Nature of Middle-Class Jobs"
Men	Women	<b>NYT, 2017:</b> "The Words Men and Women Use When They Write About Love"
Men	Women	<b>Wall Street Journal, 2016:</b> "What's Your Pay Gap?"
Men	Women	<b>DailyMail, 2018</b>
Men	Women	<b>ZEIT 2016</b>
Men	Women	<b>ZEIT 2018</b>

# Alternatives to Pink/Blue?

Some chosen alternatives?

It can backfire to invert?

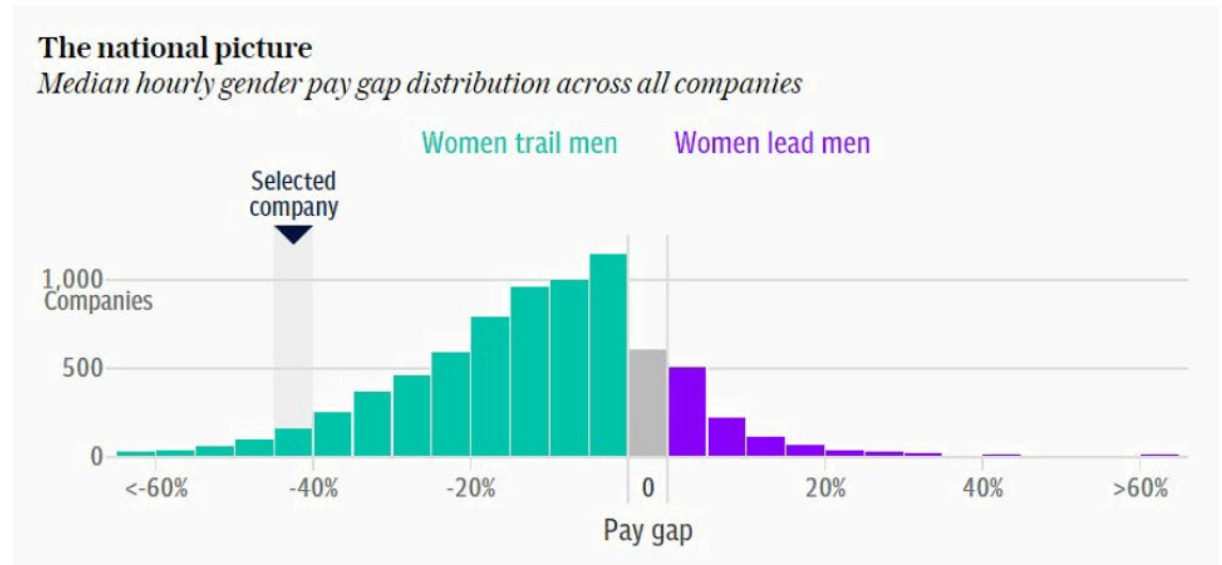


# Alternatives to Pink/Blue?

From suffrage movement

“When deciding which gender aligned with which color, it was more a case of trying to prioritize women in the order of genders. Against white, purple registers with far greater contrast and so should attract more attention when putting alongside the green, not by much but just enough to tip the scales. In a lot of the visualisations men largely outnumber women, so it was a fairly simple method of bringing them back into focus.”

Fraser Lyness, Telegraph



# Reconsider The Diverging Stacked Bar

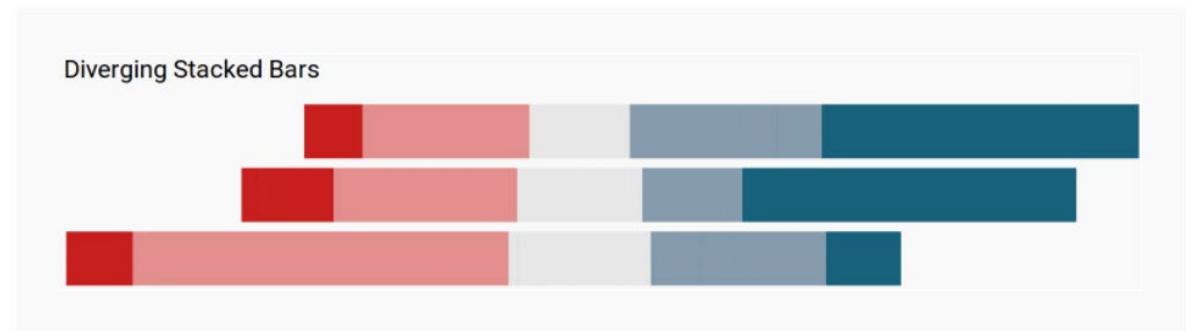
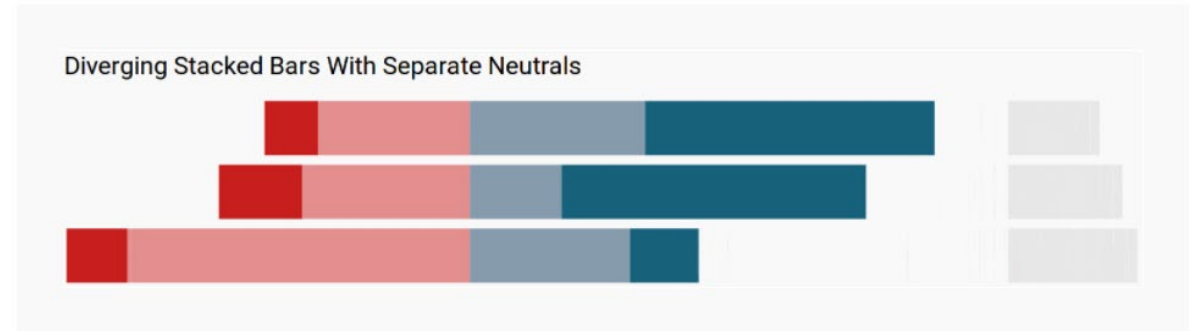
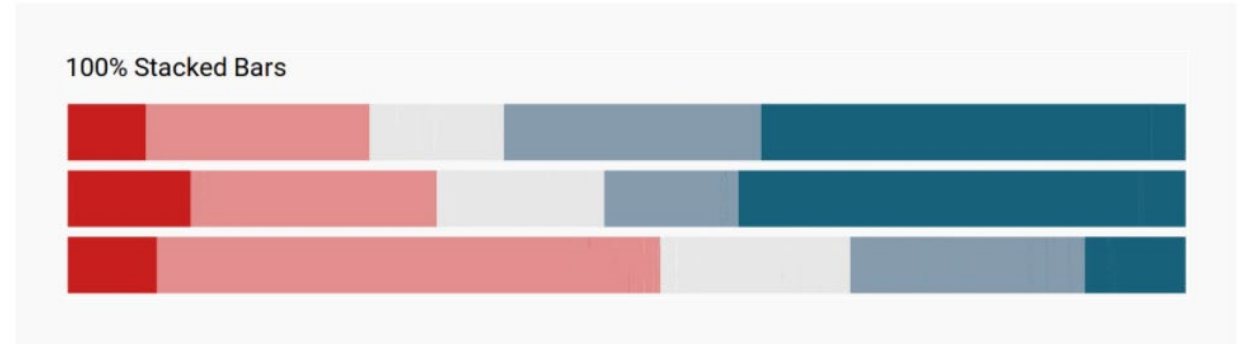
# Diverging Stacked Bar

(Common with Likert Data, percent of whole in categories)

First is best

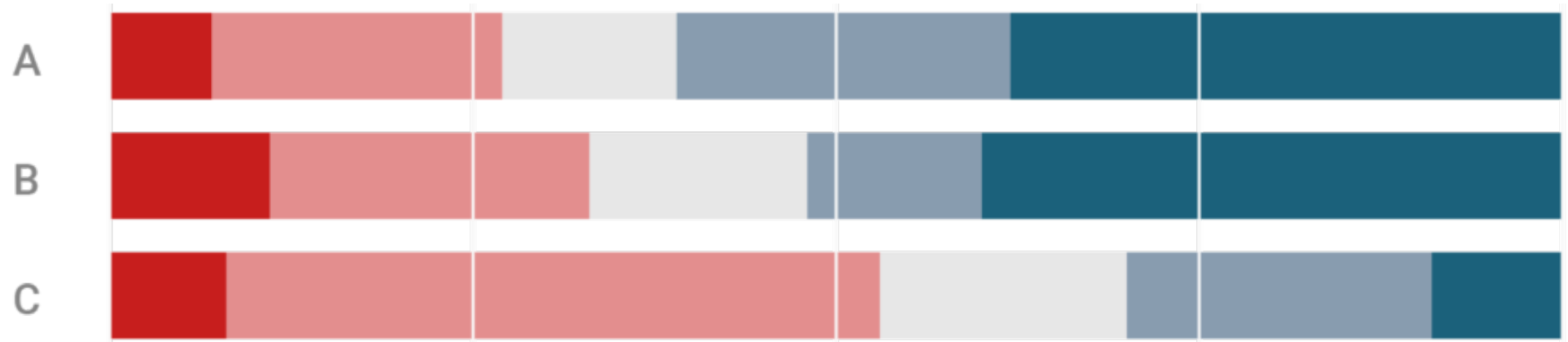
Removing neutrals is not accurate idea of data (undecideds in elections can be pivotal)

The diverging takes up more space usually, and also makes bar comparisons often harder.

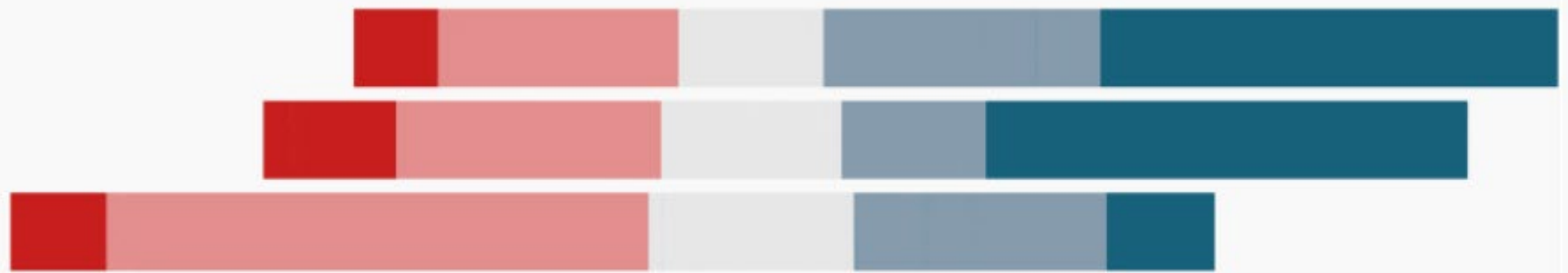


# Comparison

---



Diverging Stacked Bars



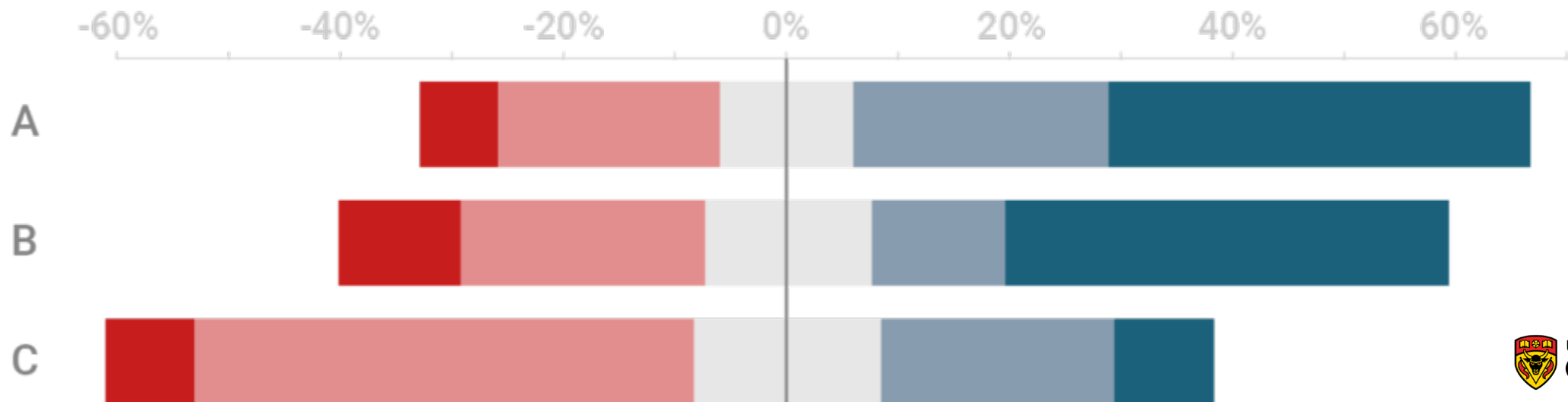


# Comparison

one of the bars in that chart share a common baseline, so none of them are comparable

Here we are primarily interested in the total percent to the right or left of the zero line;

However, even here we are misappropriating half of middle to each side

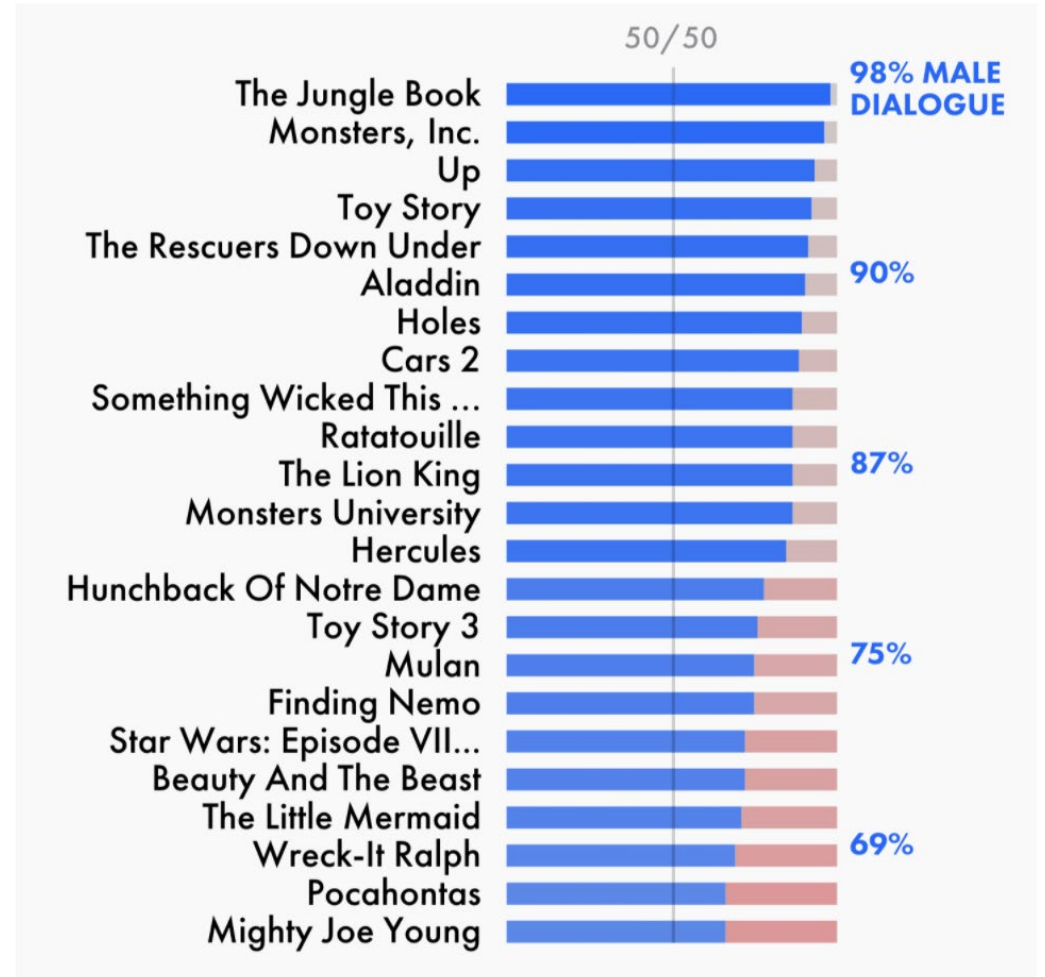
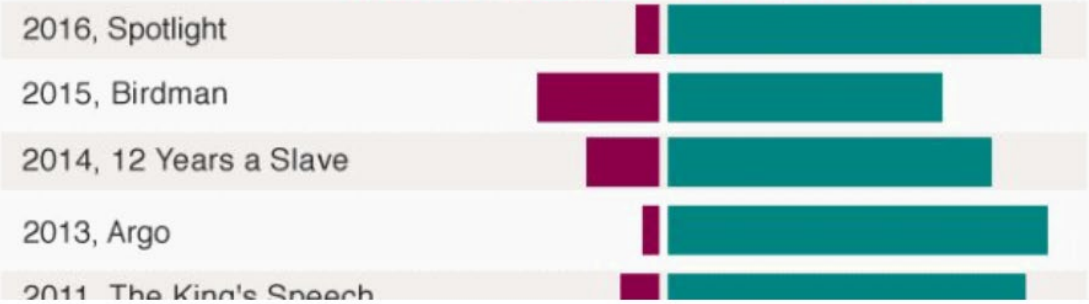


# Good example of value

## Men speak most in best picture winning films

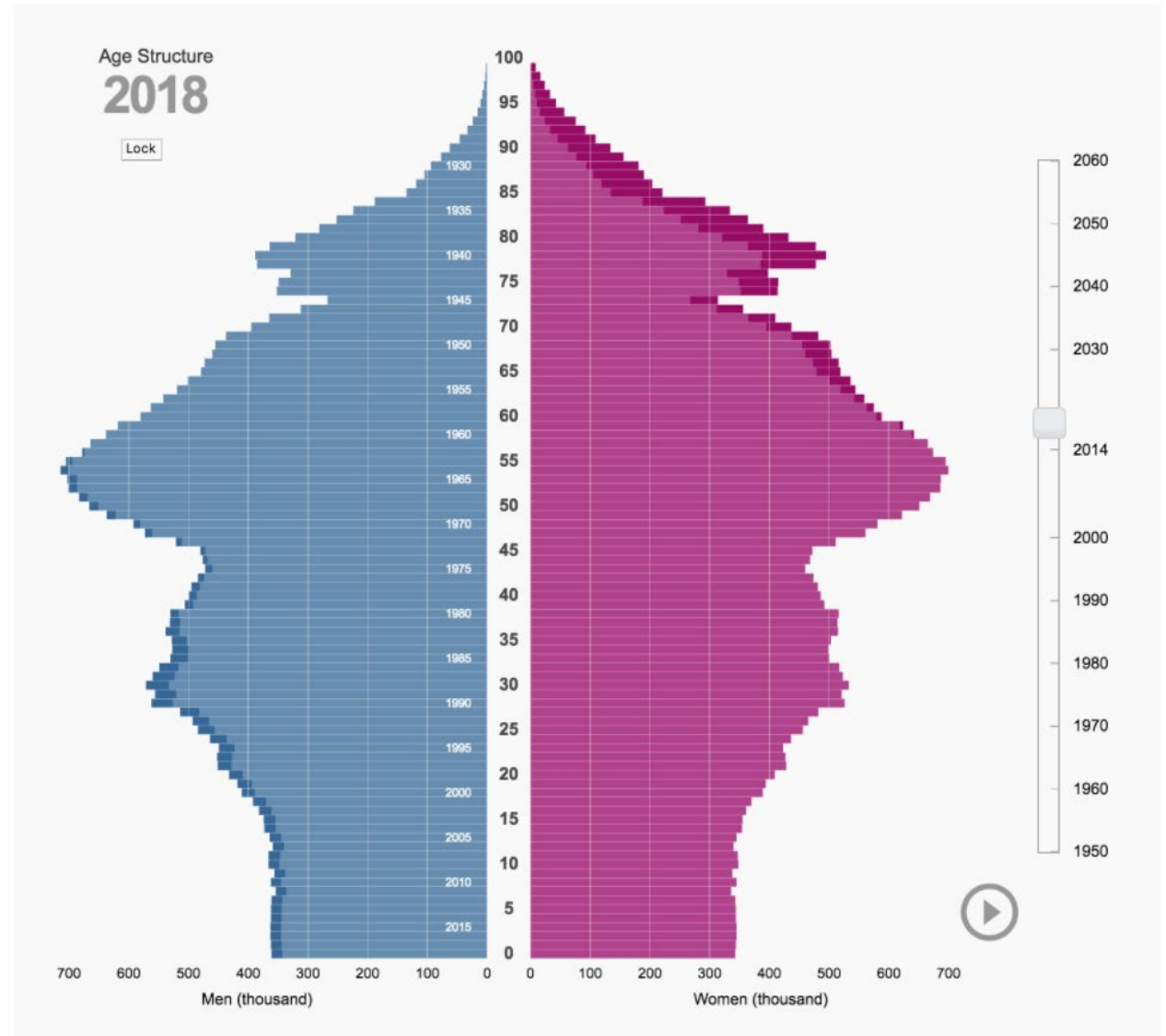
Proportion of words spoken by characters with more than 100 words

Year and film      Women speaking - Men speaking



# Not universal

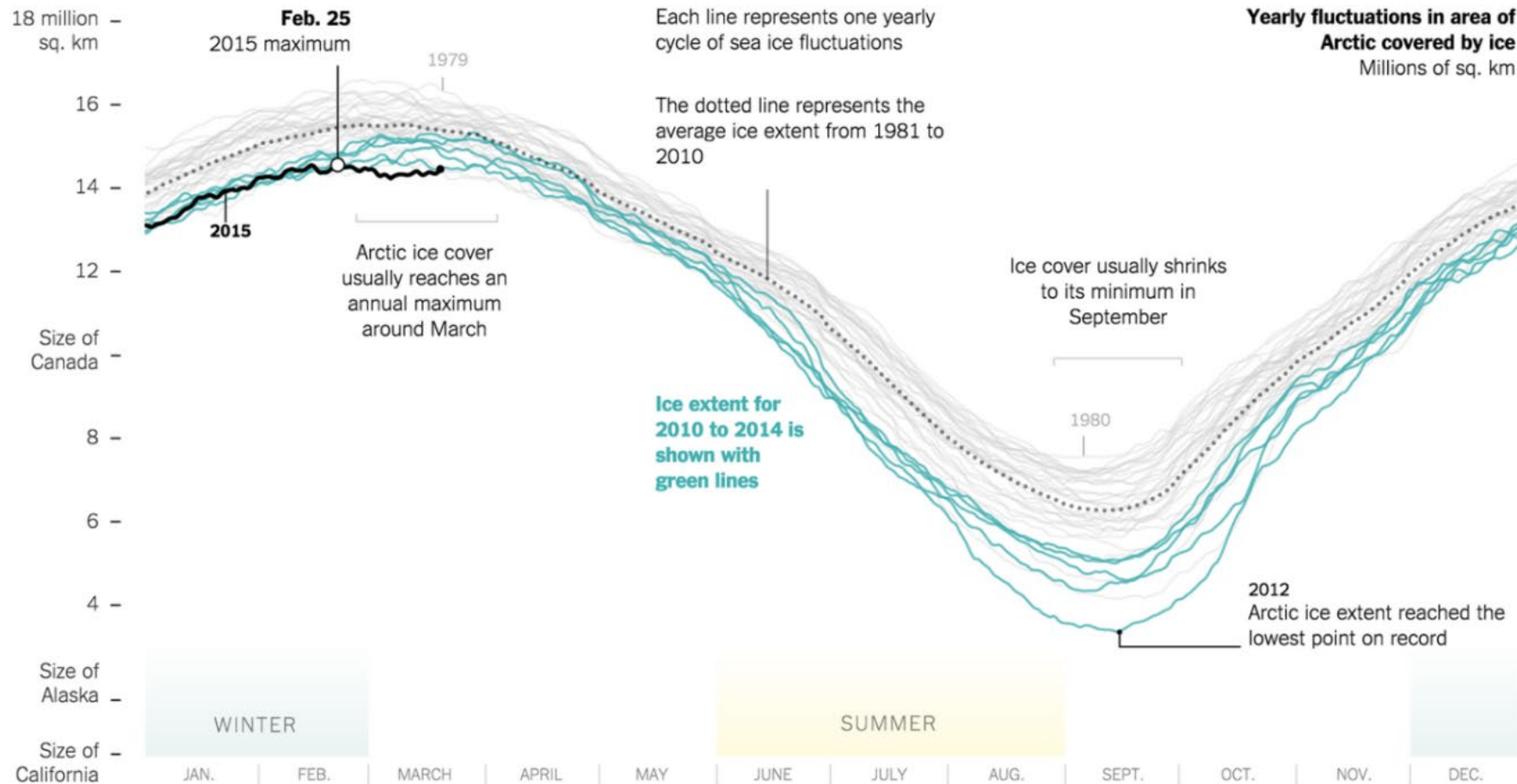
Good example where shape of relative distributions is clearer from data in divergent form



**Respect Your Reader  
(Time!)**

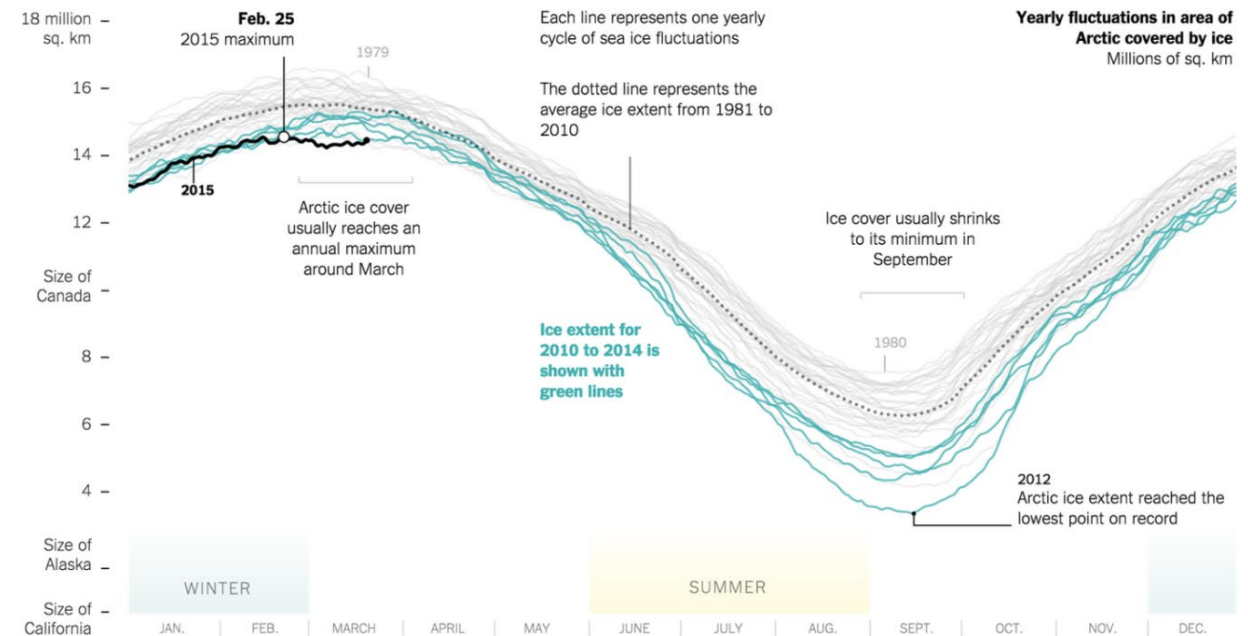
# Respect

<https://www.nytimes.com/interactive/2015/03/24/science/earth/arctic-ice-low-winter-maximum.html>



# Respect

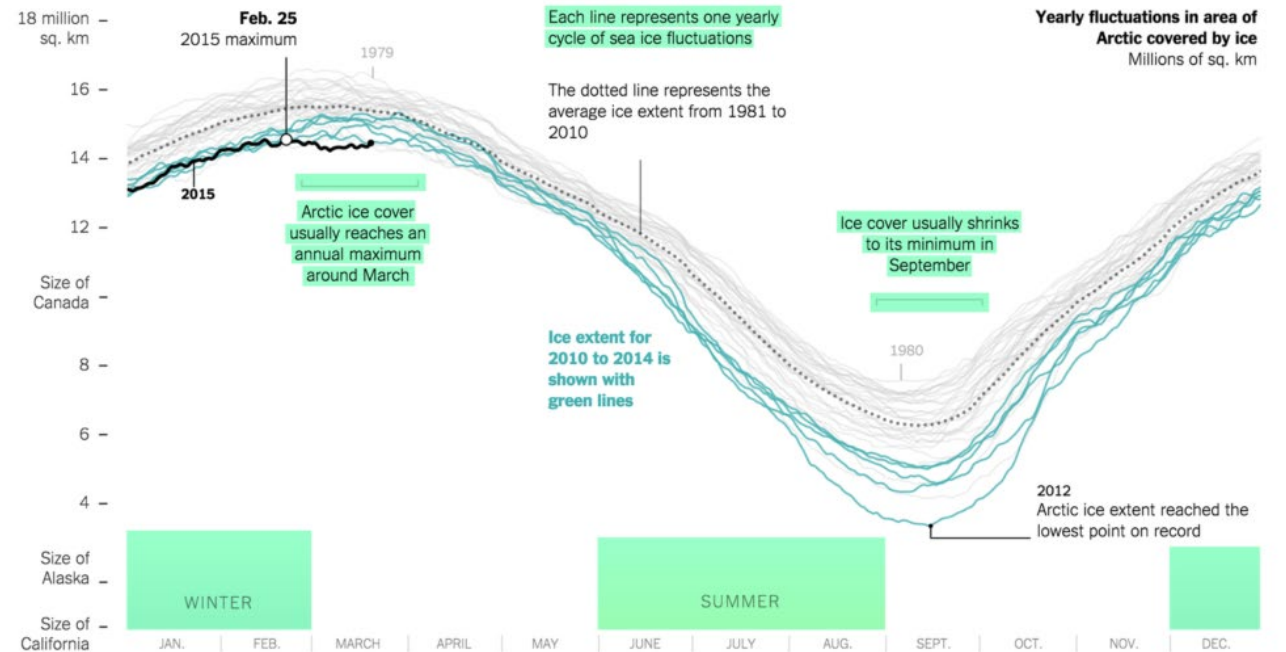
1. Whole picture is there (but subtle underneath if consider less important)
2. Orientation is clear (annotations, bold, and colouring draws to important parts)
3. Content is clear



# Respect

1. Whole picture is there (but subtle underneath if consider less important)

Reader could figure out that highs are winter, lows summer, but annotations on bottom in sections or in text nail that down (so it is not missed)

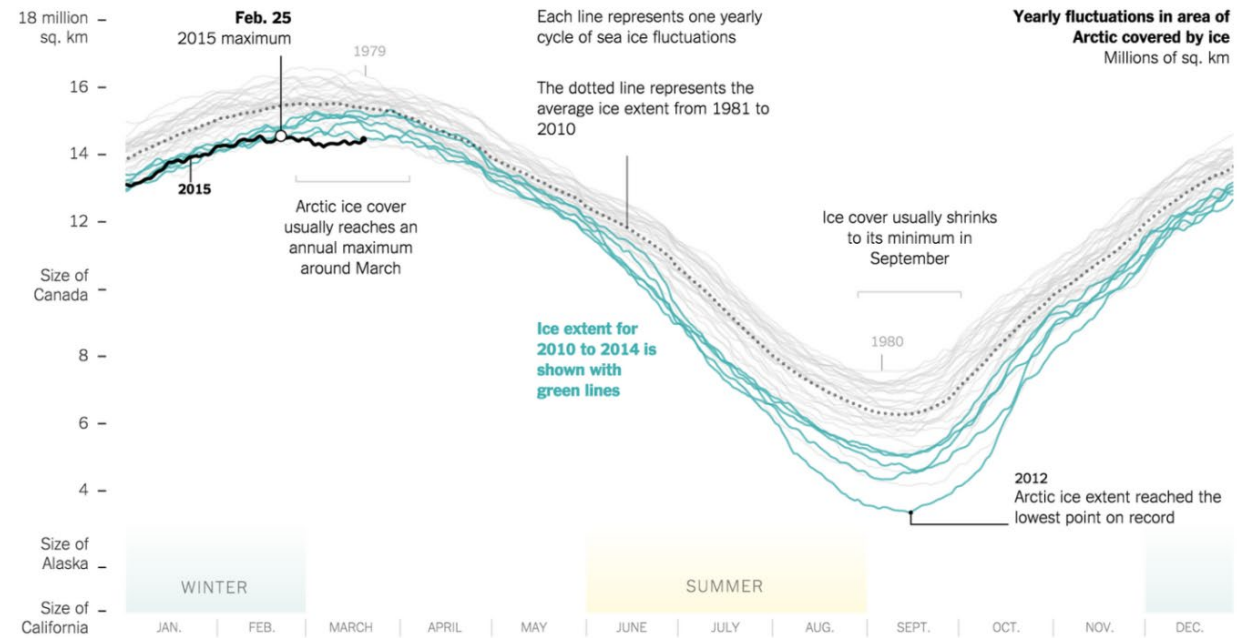
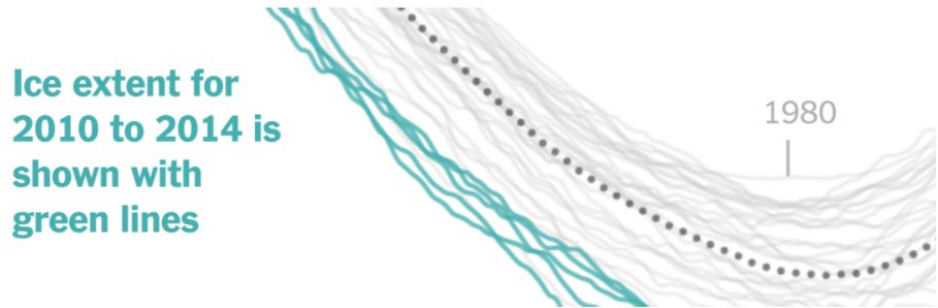


# Respect

1. Orientation is clear (annotations, bold, and colouring draws to important parts)

Eye doesn't have to search out legend for context.

Coloured annotation matches the text same colour, (in place legend).





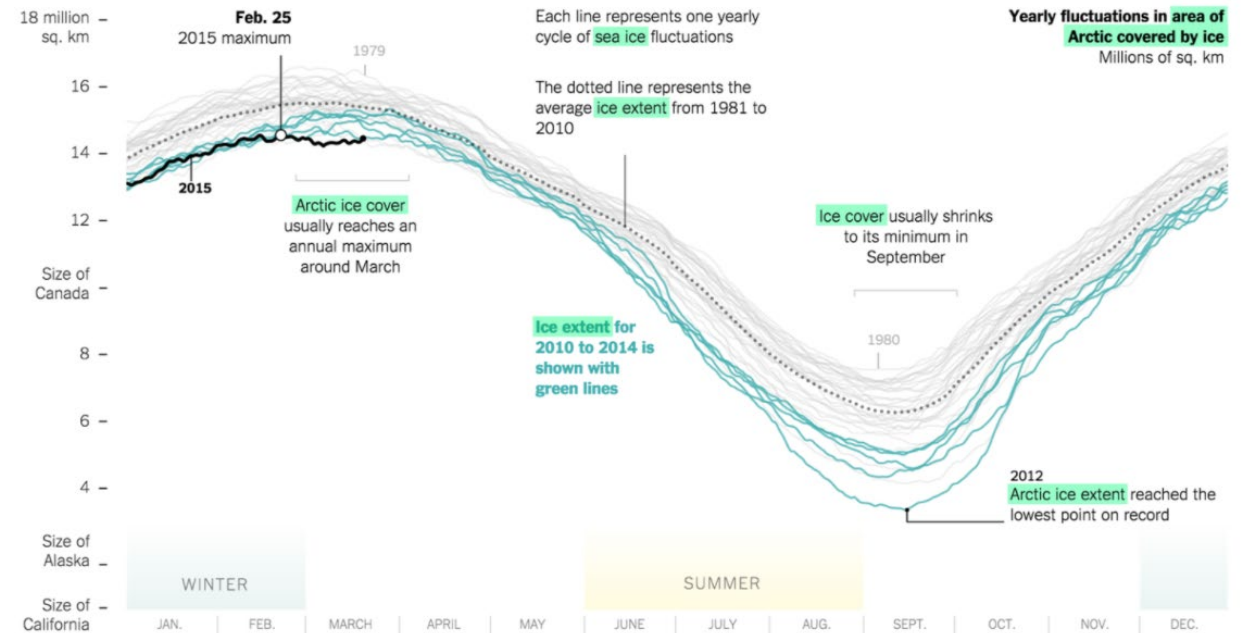
# Respect

## 1. Content is clear

### Redundancy

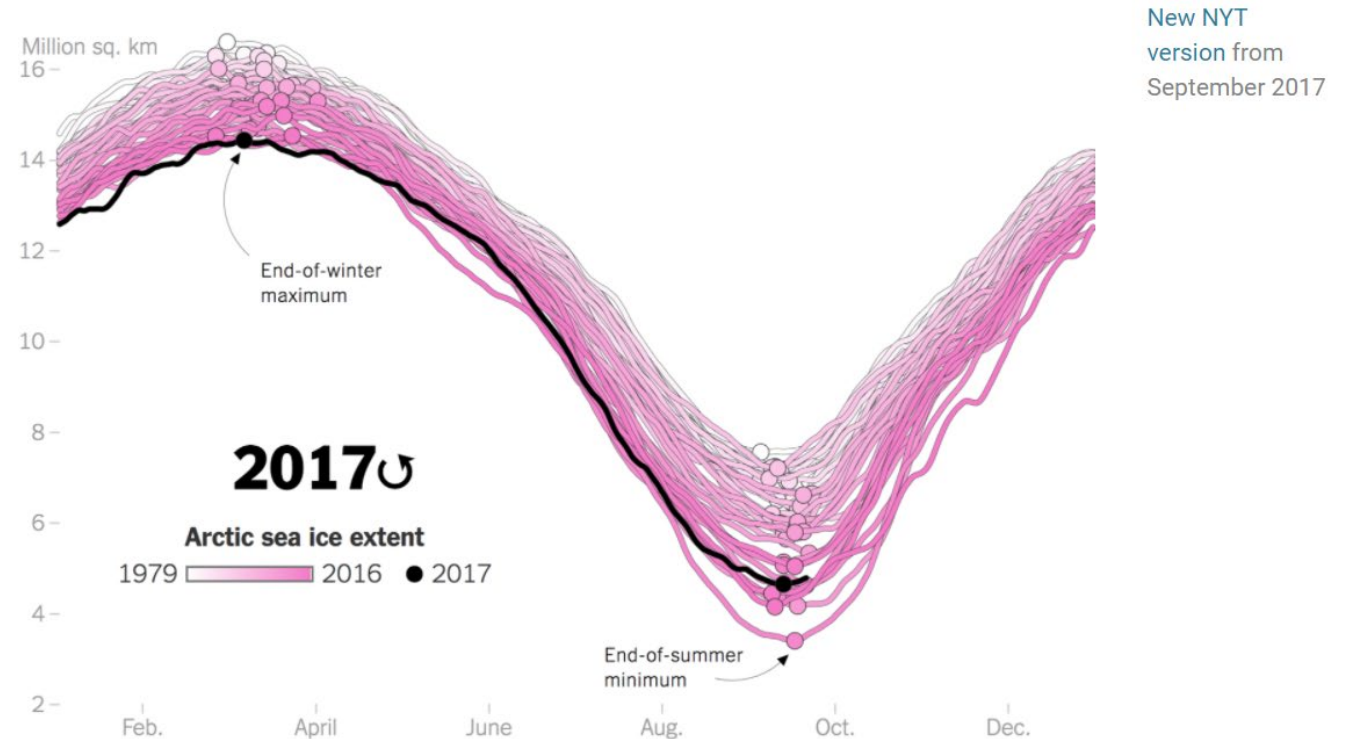
Don't rely on reader to chase the idea. Look for axis, axis titles, legend, chart titles.

They matter and should be there, but we see info repeated where it is most useful.



# The Dearth of Mobile Reading

1. Annotations lost in size
2. Changed to interactive version (popular right now but a lot of unwanted consequences)



# Some Current Examples

# Ukraine (2022)

---

- <https://twitter.com/LeviWesterveld/status/1506253203704356867>
- Evolution over time NYTimes
- Gives appearances of strong success



Source: New York Times reporting; Institute for the Study of War (Russian-occupied areas). Data as of 1 p.m. Eastern on Feb. 26. • The New York Times

# Ukraine (2022)

- <https://twitter.com/LeviWesterveld/status/1506253203704356867>

- Evolution over time NYTimes
- Smaller arrows, shows weakness areas



Source: Institute for the Study of War with American Enterprise Institute's Critical Threats Project (Russian-occupied areas) | Note: Occupied areas are as of 3 p.m. Eastern on March 3. • The New York Times

# Ukraine (2022)

- <https://twitter.com/LeviWesterveld/status/1506253203704356867>

- Evolution over time NYTimes
- Shows opponent, and defense



Source: Institute for the Study of War with American Enterprise Institute's Critical Threats Project (Russian-occupied areas) • Note: Occupied areas are as of 3 p.m. Eastern on March 7. Russian forces in eastern Ukraine include Russian-backed separatists. • The New York Times

# Ukraine (2022)

- <https://twitter.com/LeviWesterveld/status/1506253203704356867>
- Evolution over time NYTimes
- Shows occupational of cities, don't 'give' areas to invader by dropping red areas



Source: Verified imagery, witness accounts, Ukrainian and Russian officials, Institute for the Study of War with American Enterprise Institute's Critical Threats Project • Note: Russian forces in eastern Ukraine include Russian-backed separatists. Approximate areas of fighting are drawn from reports from March 13 to 21. • The New York Times

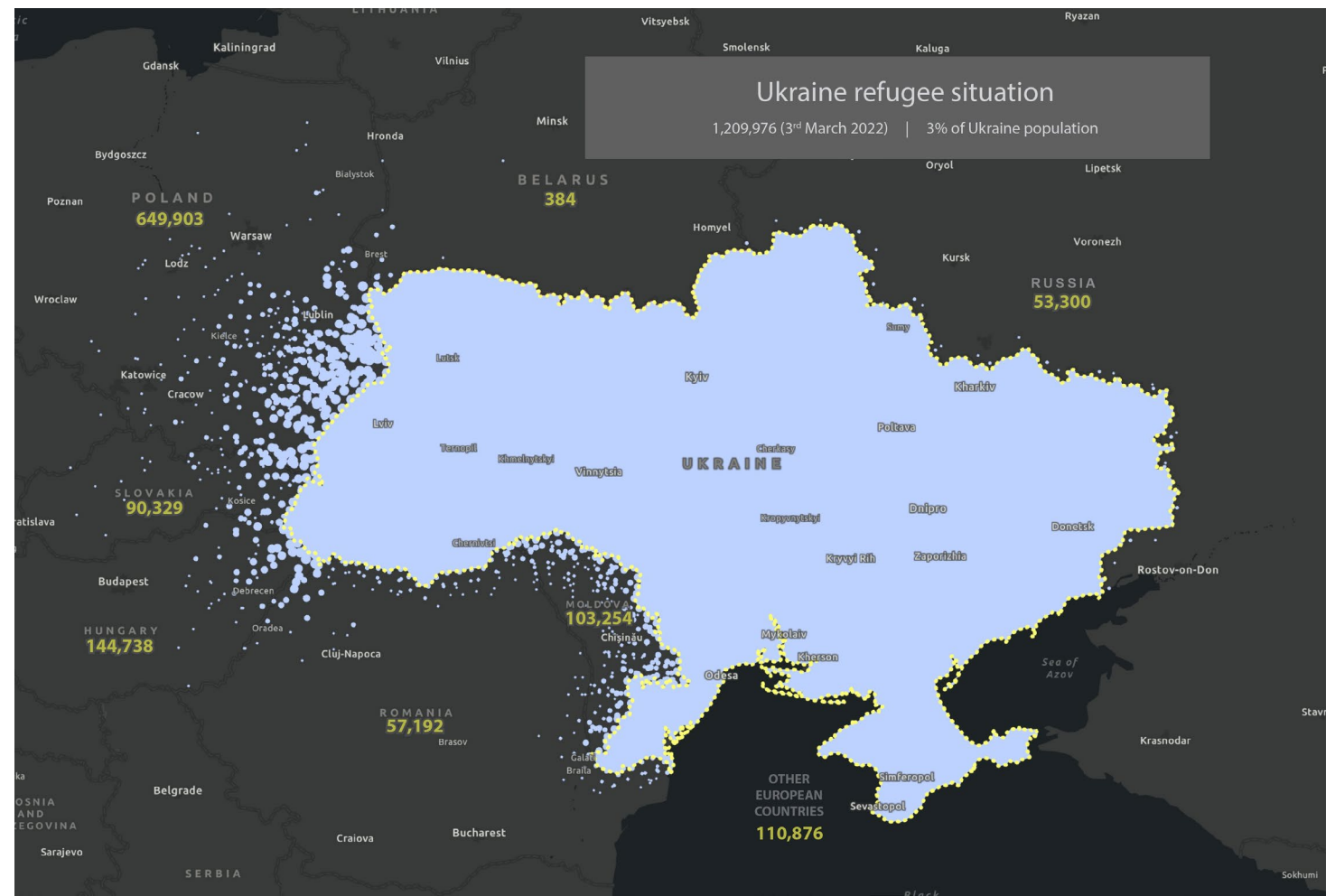
# Ukraine (2022)





# Ukraine (2022)

- <https://twitter.com/kennethfield/status/1499907392338751488/photo/1>



# Ukraine (2022)

---

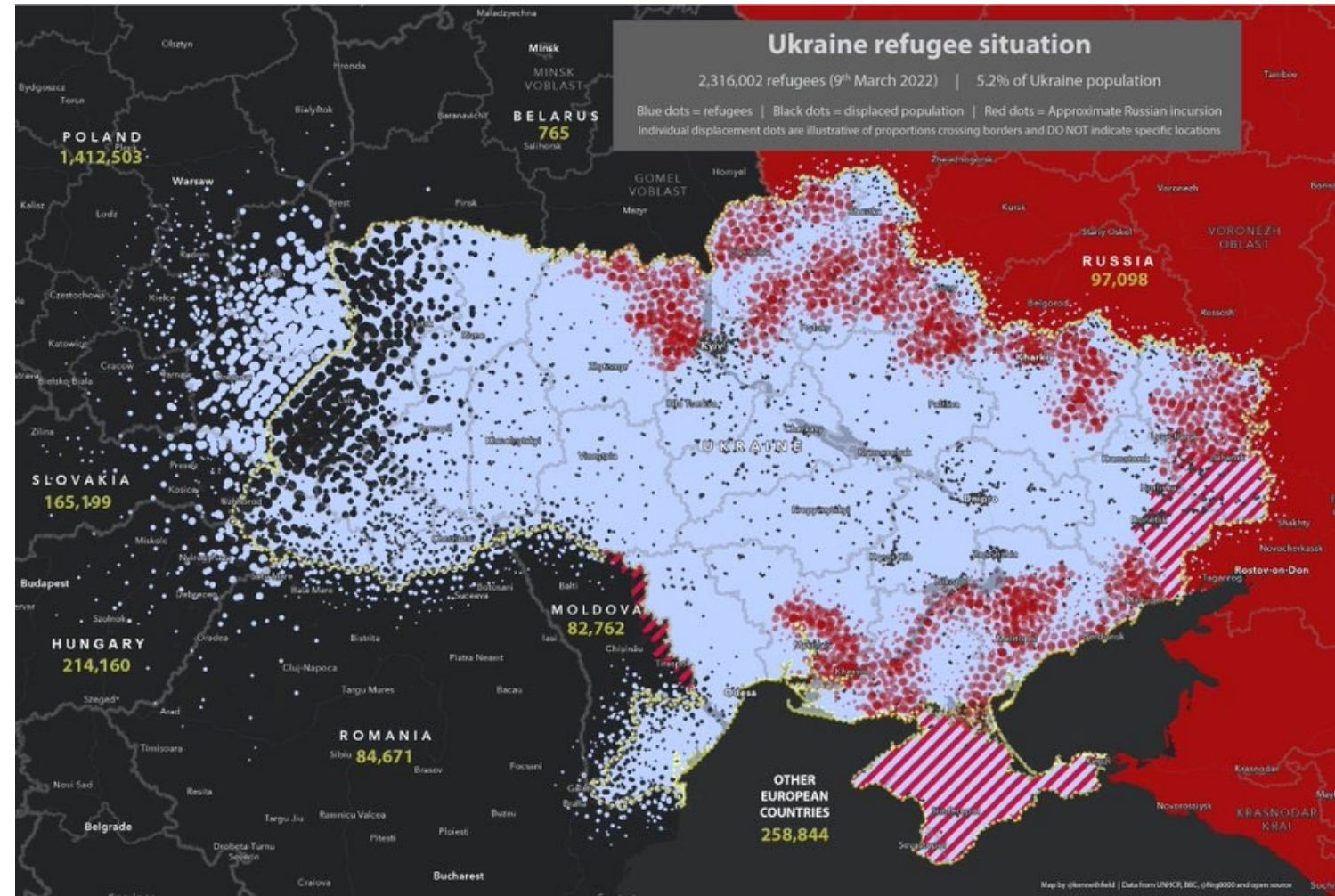
- <https://twitter.com/kennethfield/status/1500235419966849028/photo/1>
- Negative space



# Ukraine (2022)

- <https://twitter.com/kennethfield/status/1502083106370555907>

- Negative space



# Kenneth Field

- <https://twitter.com/kennethfield/status/1499907392338751488/photo/1>
- Kenneth Field
- @kennethfield
- Cartography is my passion & profession. Author of **CARTOGRAPHY. & THEMATIC MAPPING**
- Example different choices for election presentation

Thematic maps of the 2016 Presidential election (lower 48 states)



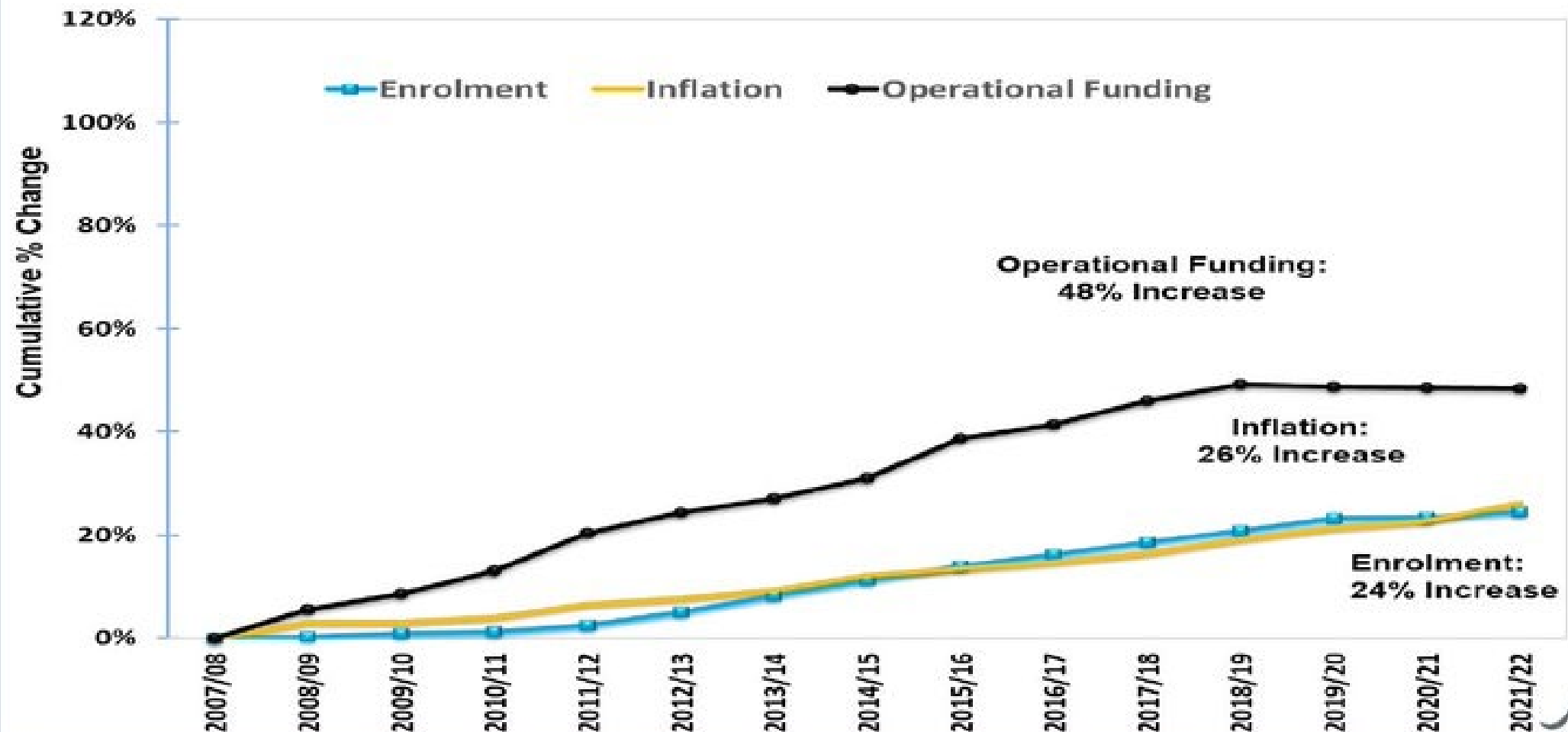
# Alberta Education Funding

- <https://twitter.com/AdrianaLaGrange/status/1503835502973124609>
- Not post-secondary conversation (not our budget), not that cuts have not occurred in PSE



# Student Enrolment Growth – A Historical Perspective

Student Enrolment Growth – A Historical Perspective (2007- 2022)



# Alberta Education Funding

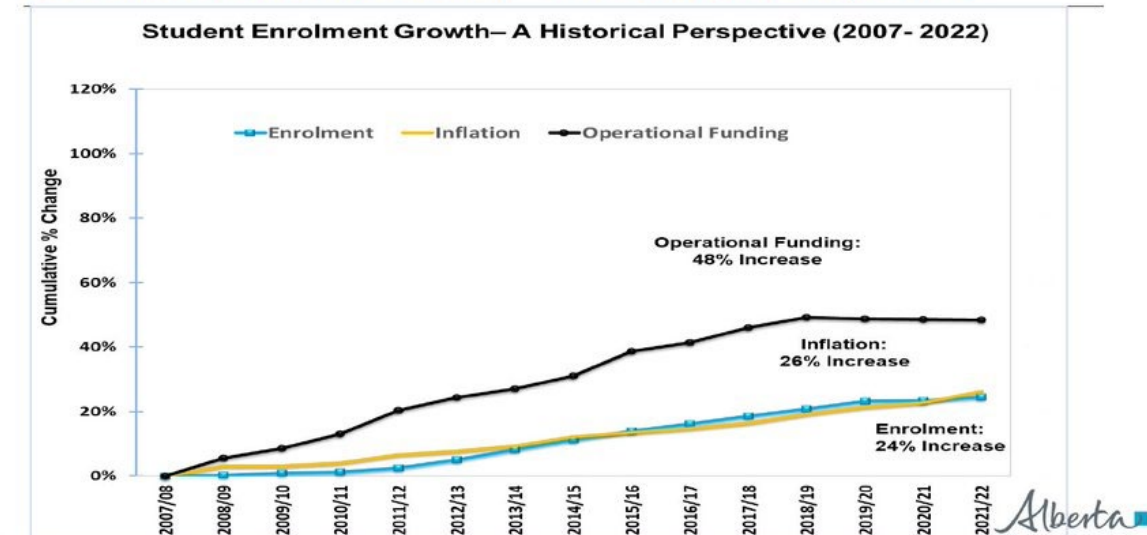


Lee Francis  
@yegronin

In the words of my teenagers “OMG! I can’t even!”.

- Per capita , inflation is ignored

## Student Enrolment Growth – A Historical Perspective



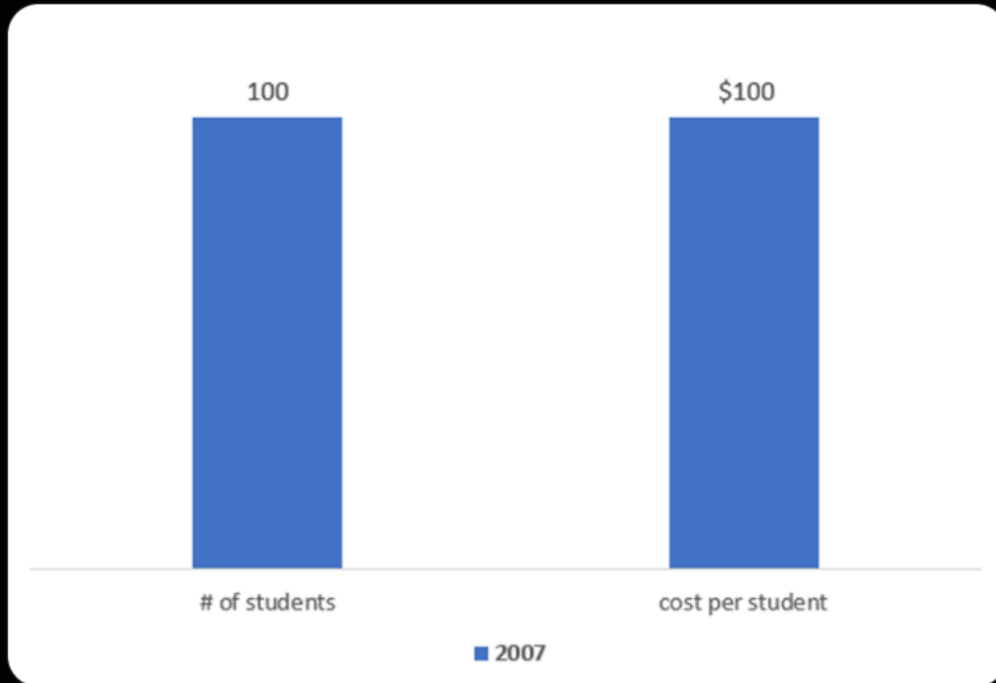
# Alberta Education Funding



Lee Francis @yegronin · Mar 15

Replying to @yegronin

All I have to work with is the minister's % increases. It's enough to prove my point. We start in 2007. Suppose we have 100 students and the cost of educating each student is \$100. 2/6

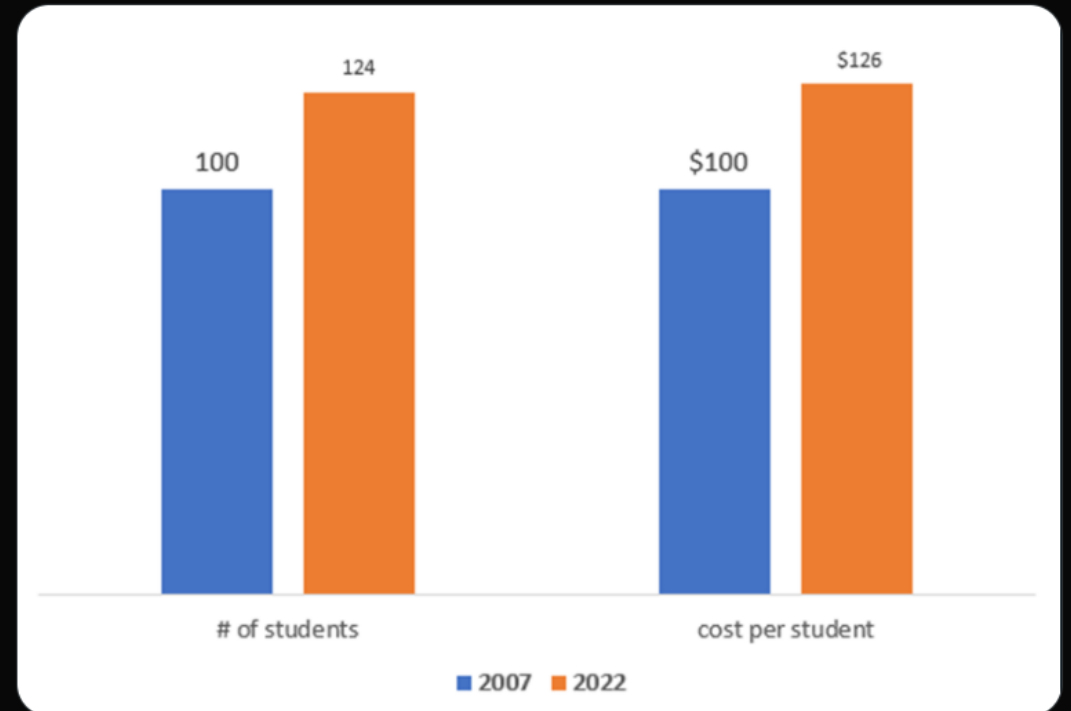


1 8 59



Lee Francis @yegronin · Mar 15

Now let's fast-forward to 2022. Enrolment has grown 24% so we now have 124 students. Inflation grew 26% so the cost of educating each student is now \$126/student. 3/6



1 7 60

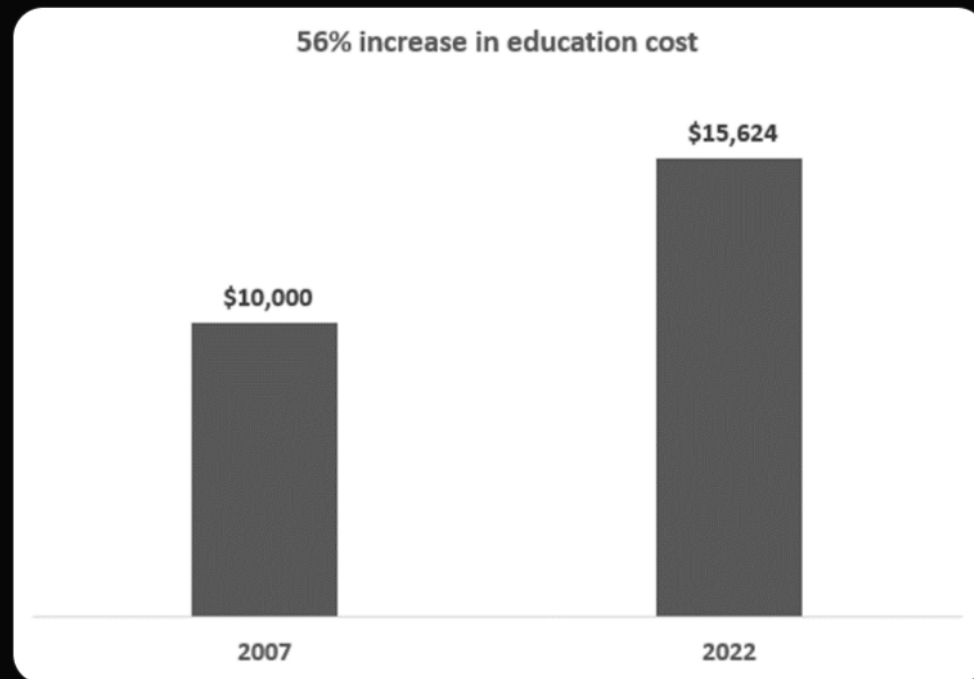


# Alberta Education Funding



Lee Francis @yegronin · Mar 15

Still with me? Great! So how much does it cost to educate all students in 2007 & 2022? In 2007 it cost \$10,000 (100 students x \$100/student). In 2022 it cost \$15,624 (124 students x \$126/student). That's a 56% increase. 4/6

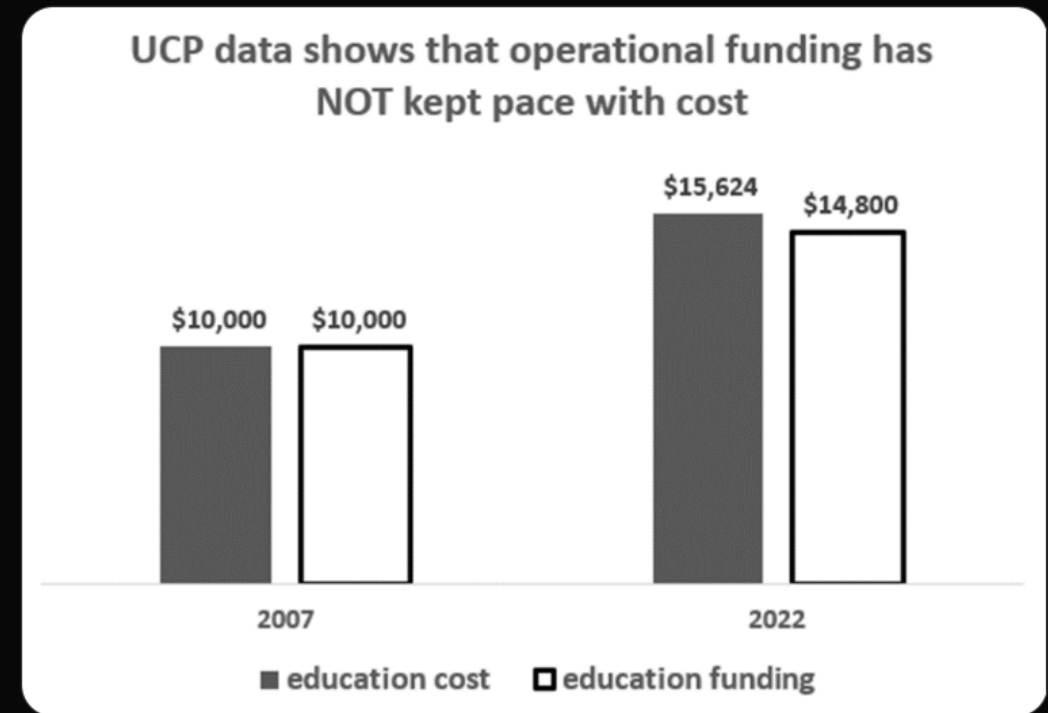


1 7 65



Lee Francis @yegronin · Mar 15

On to operational funding. Let's assume that we had just enough funding to cover costs in 2007 and make it \$10,000. By 2022 that funding increases 48% which brings it to \$14,800. 5/6



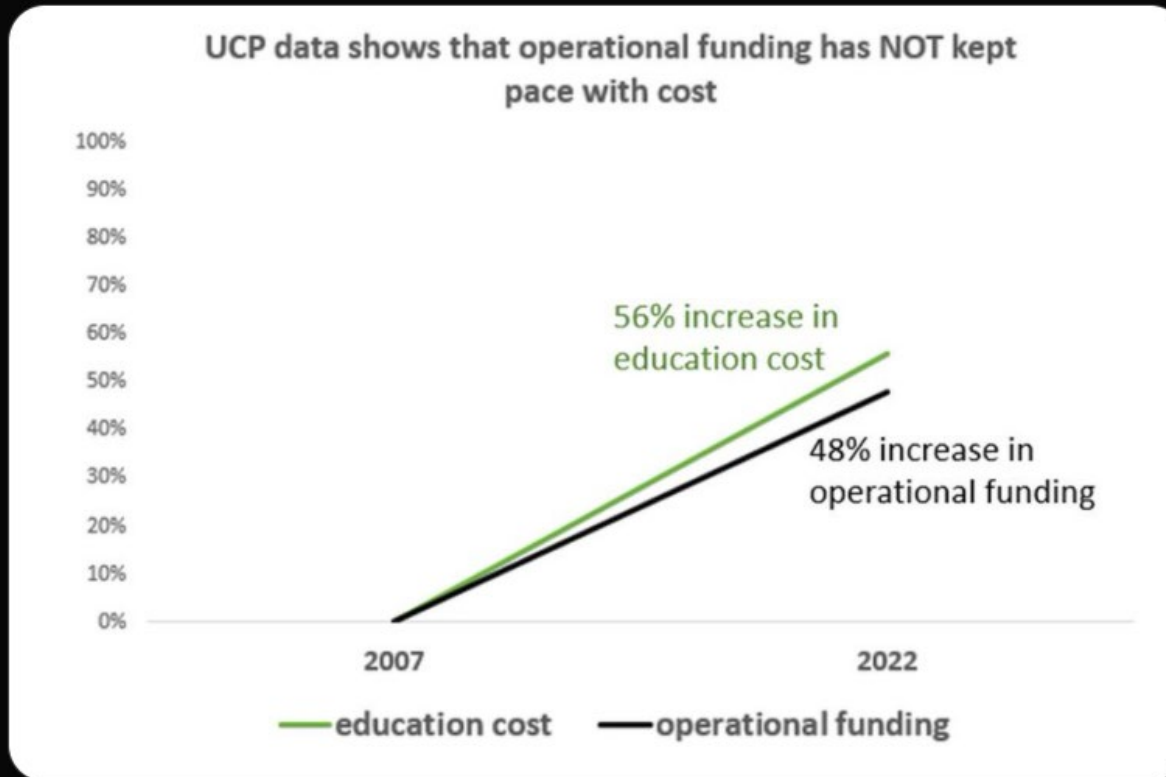
1 6 59

# Alberta Education Funding



Lee Francis @yegronin · Mar 15

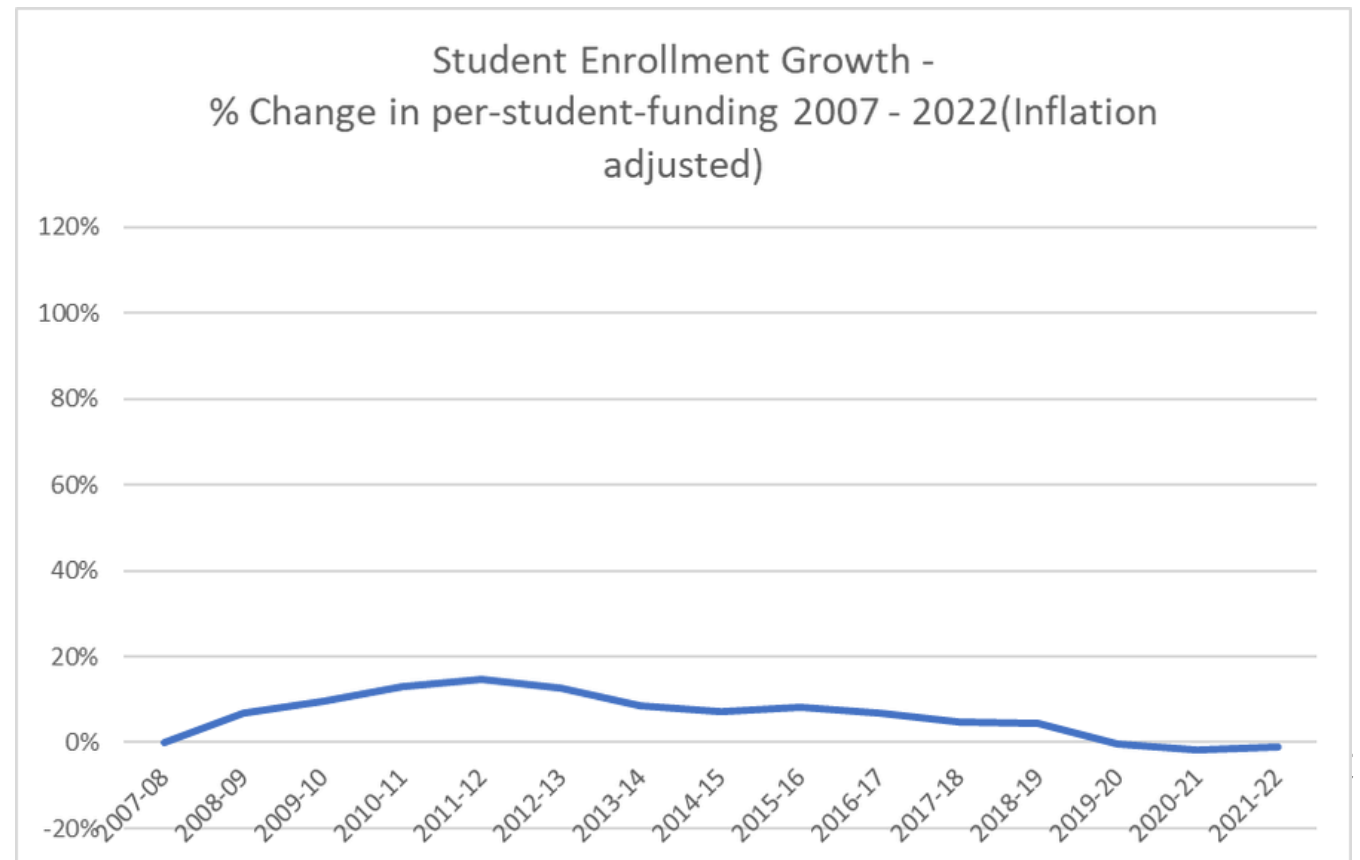
Would you prefer a line graph? Fine. Here you go. Yeesh! 6/6



# Alberta Education Funding

---

- <https://twitter.com/SeanDunn10/status/1503886553445064704>
- Relative growths instead? (this doesn't even include things like inflation)



# Onward to ... DATA 211?

---

Jonathan Hudson  
[jwhudson@ucalgary.ca](mailto:jwhudson@ucalgary.ca)  
<https://pages.cpsc.ucalgary.ca/~jwhudson/>



UNIVERSITY OF  
CALGARY