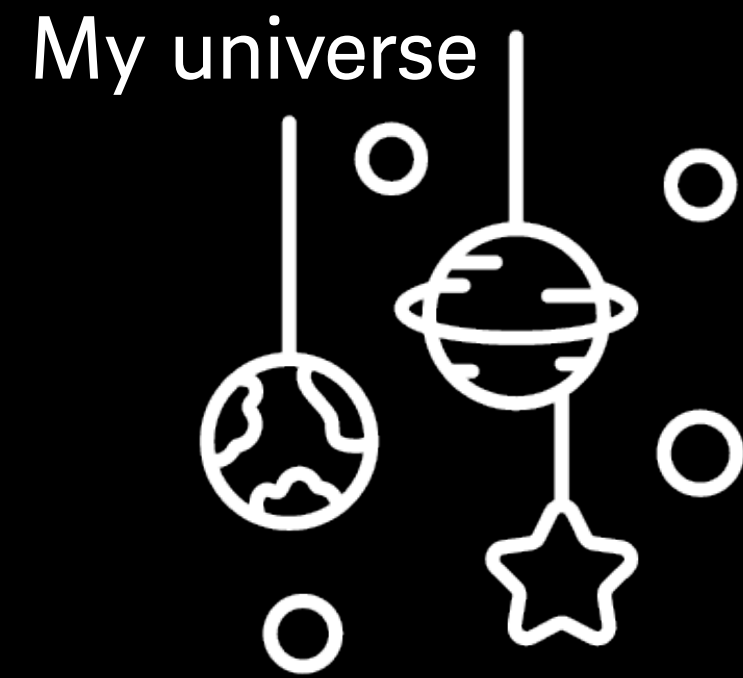


Bringing the power of dataviz to everyone

Coda.Br
November 10, 2021

Gurman Bhatia
@gurmanbhatia
www.gurmanbhatia.com

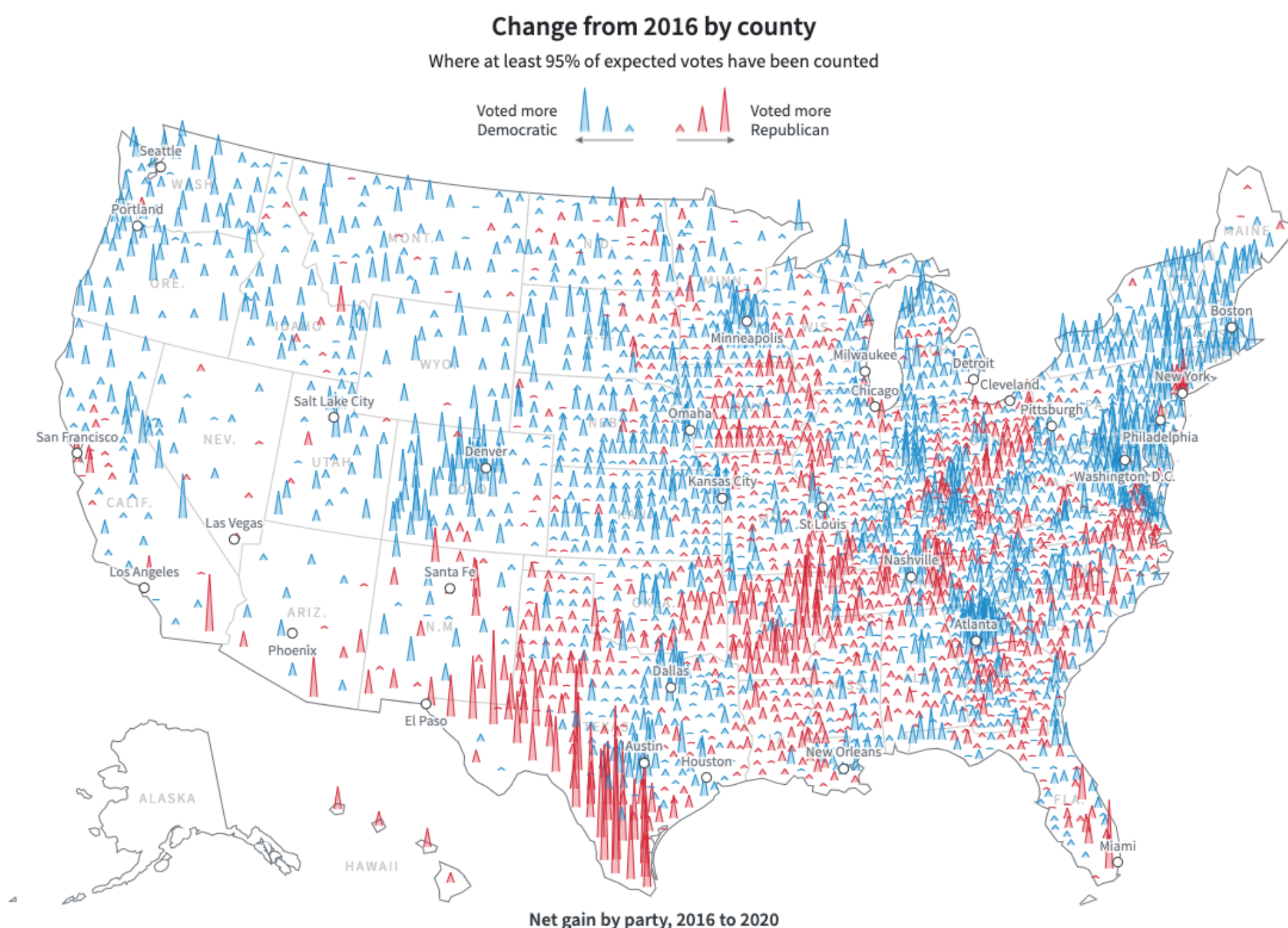


Journalism Teaching Dataviz Code

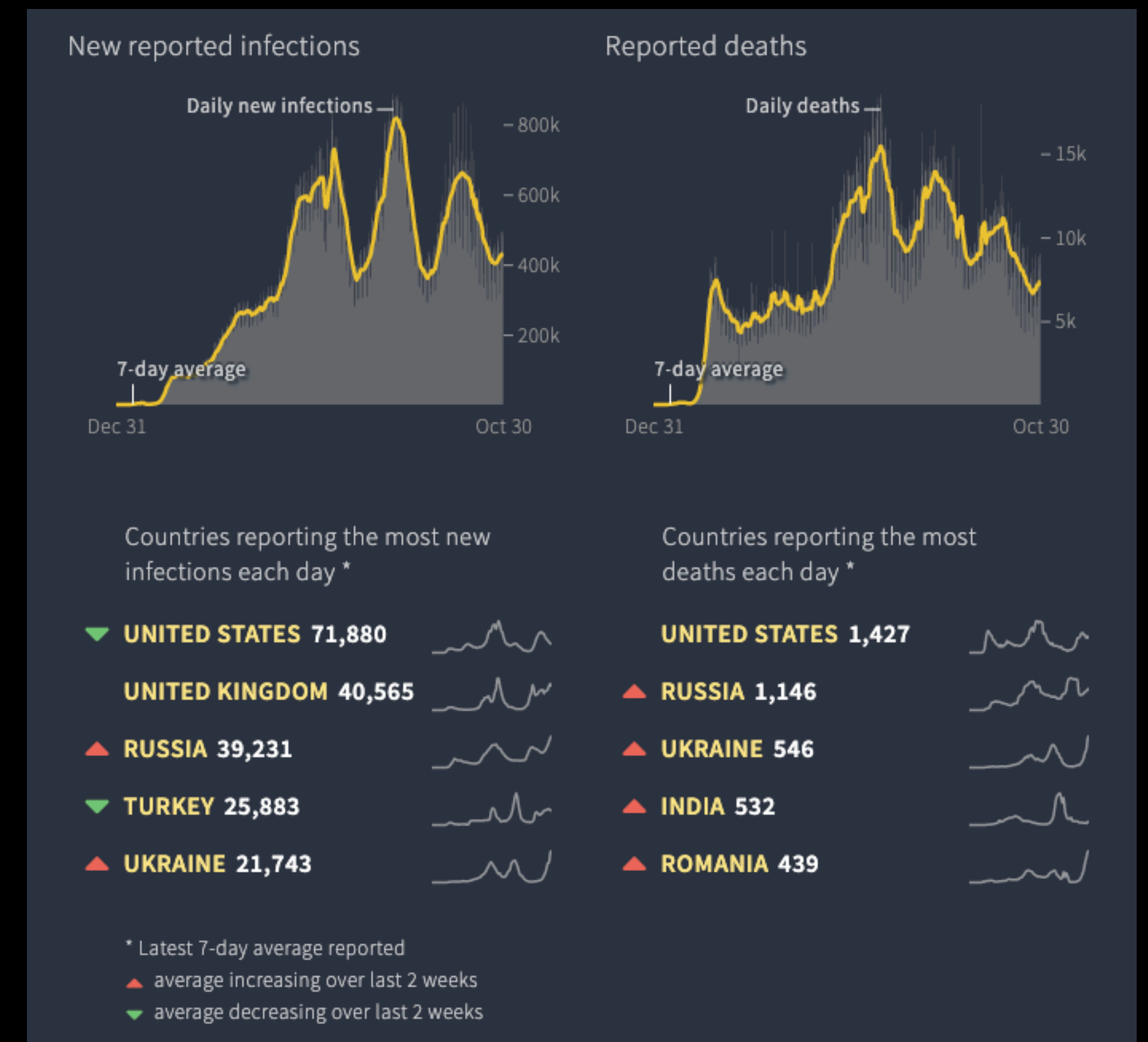
**The dual
worlds of**

**data
visualisation**

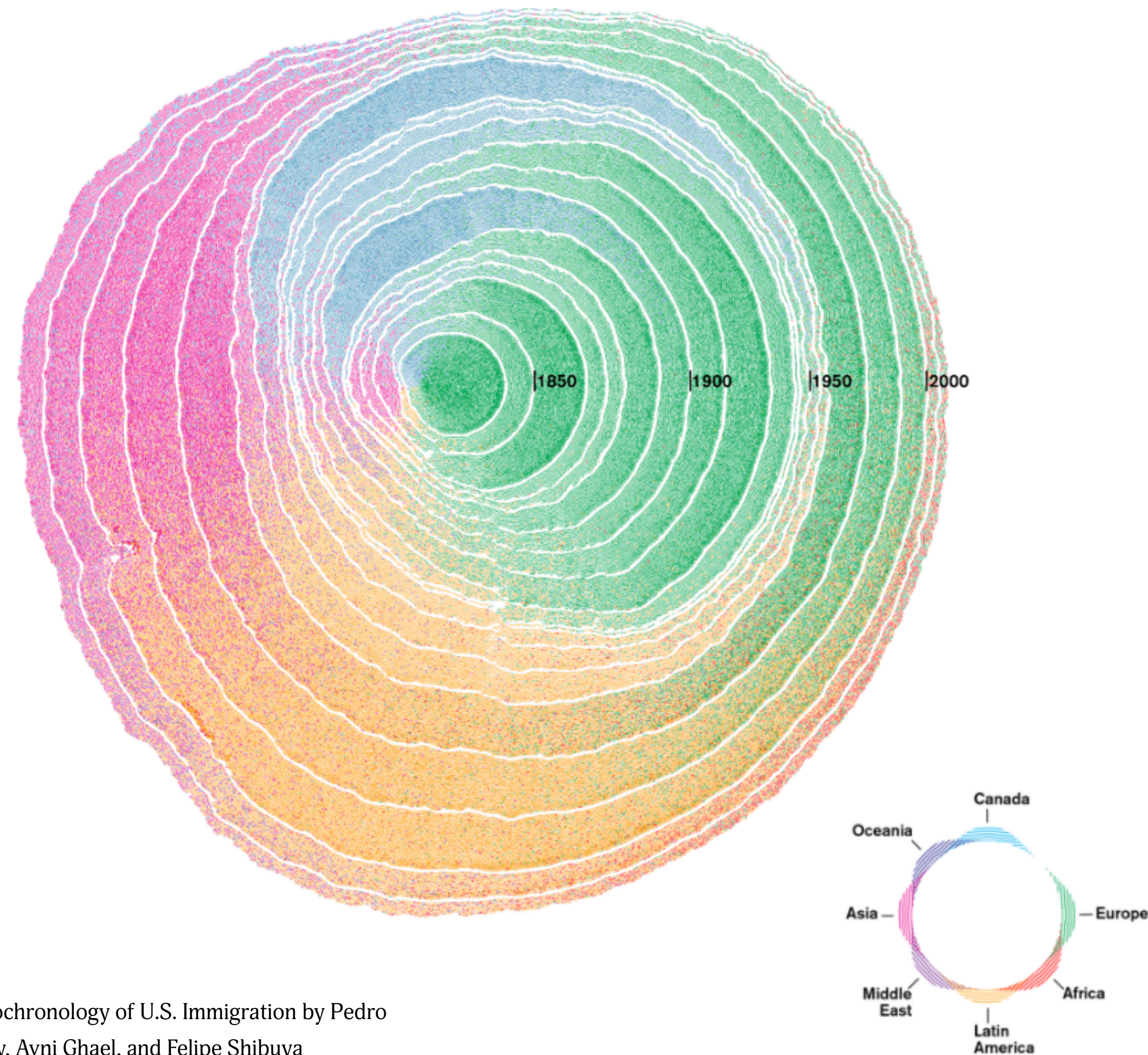
Stories & explanations



Dashboards & explorations



Bespoke and customised

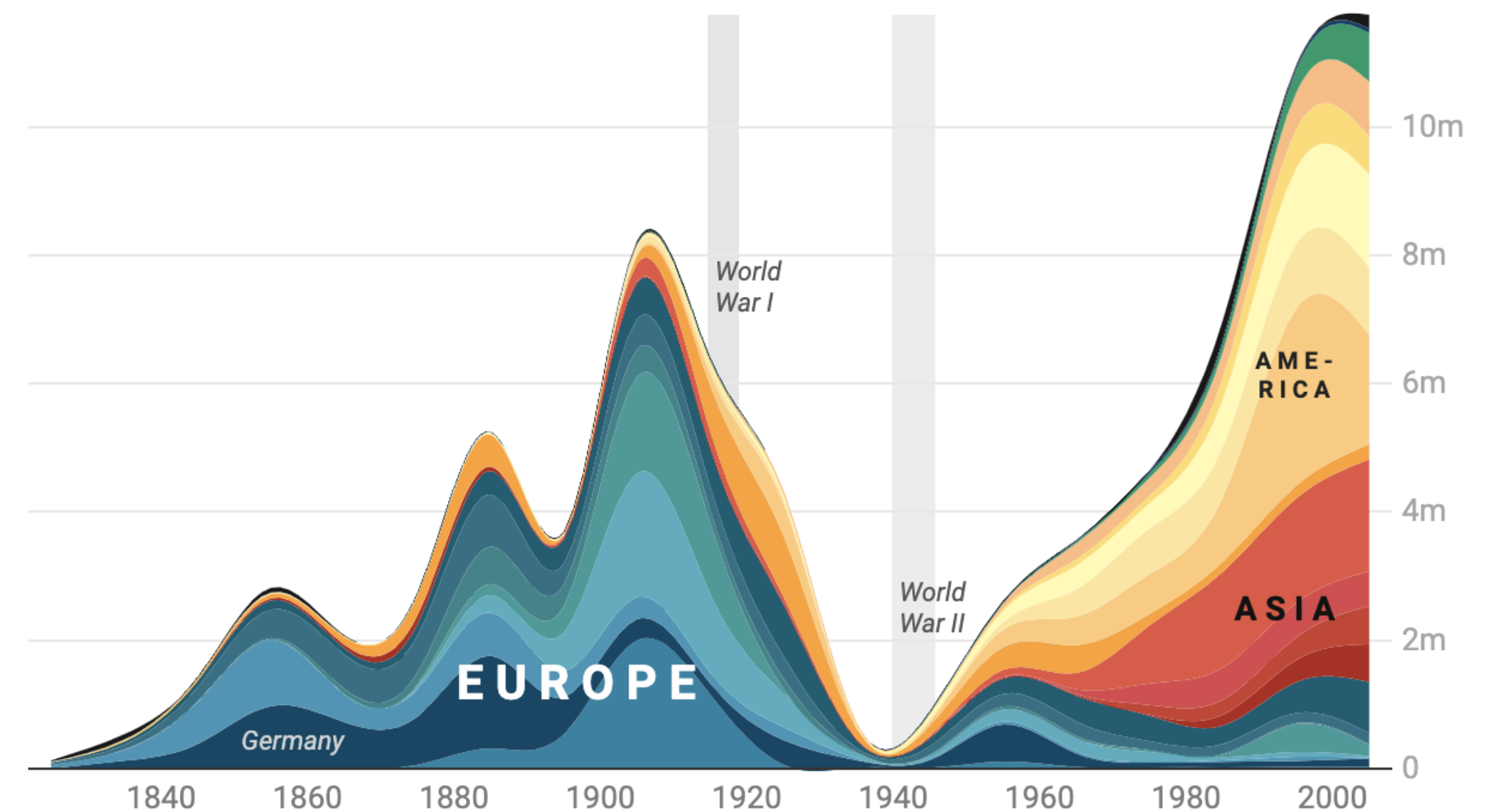


Simulated Dendrochronology of U.S. Immigration by Pedro Cruz, John Wihbey, Avni Ghael, and Felipe Shibuya

Traditional and tool-driven

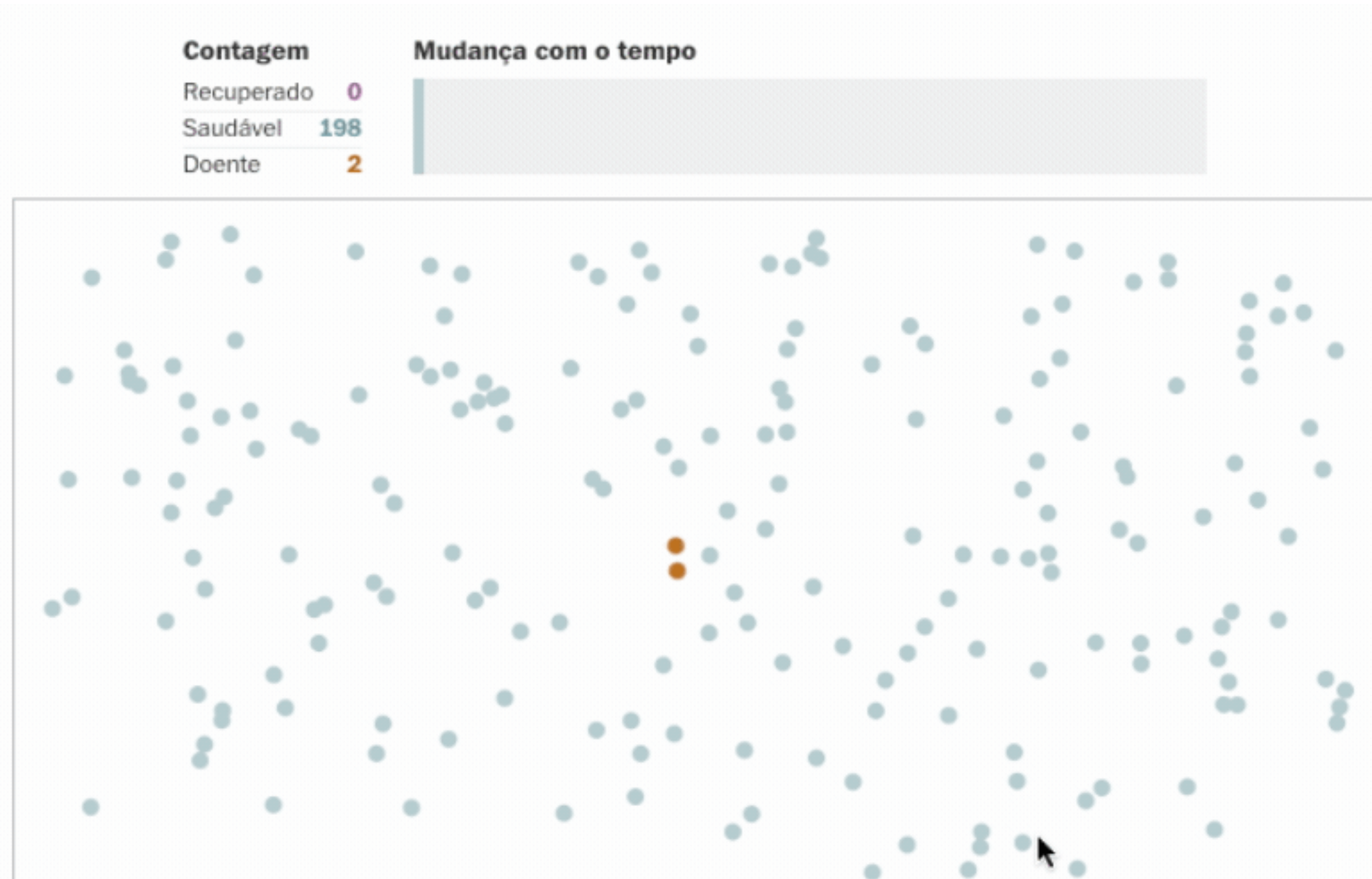
Migration to the US by world region, 1820-2009

The numbers are recorded by decade. For example, the numbers recorded for 1905 tell us the number of immigrants between 1900 and 1910.



www.gurmanbhatia.com

Interactive



Por que surtos como o coronavírus se espalham exponencialmente e como "achatar a curva", Washington Post

Static

NATIONAL PARTIES

K Surendran
240 cases



Bharatiya Janata Party

Dean Kuriakose
201 cases



Indian National Congress

Yogesh Verma
27 cases



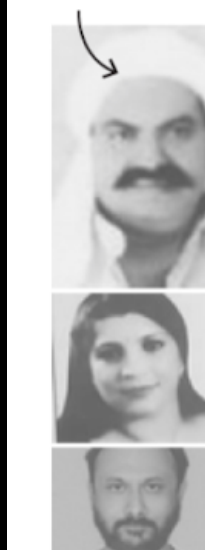
BSP

Others

CPI (M)

INDEPENDENTS

Ateek Ahmad - 54 cases



REGIONAL PARTIES

Oli Mohammad Mallick - 32 cases



Shiv Sena

Others

Samajwadi

Many forms

Many roles

Stories
Explanatory
Bespoke
Customised
Interactive

Dashboards
Exploration
Traditional
Tool-driven
Static

All these forms
are gaining
popularity



Data Journalism



Source: Reuters Graphics
Vaccination Tracker

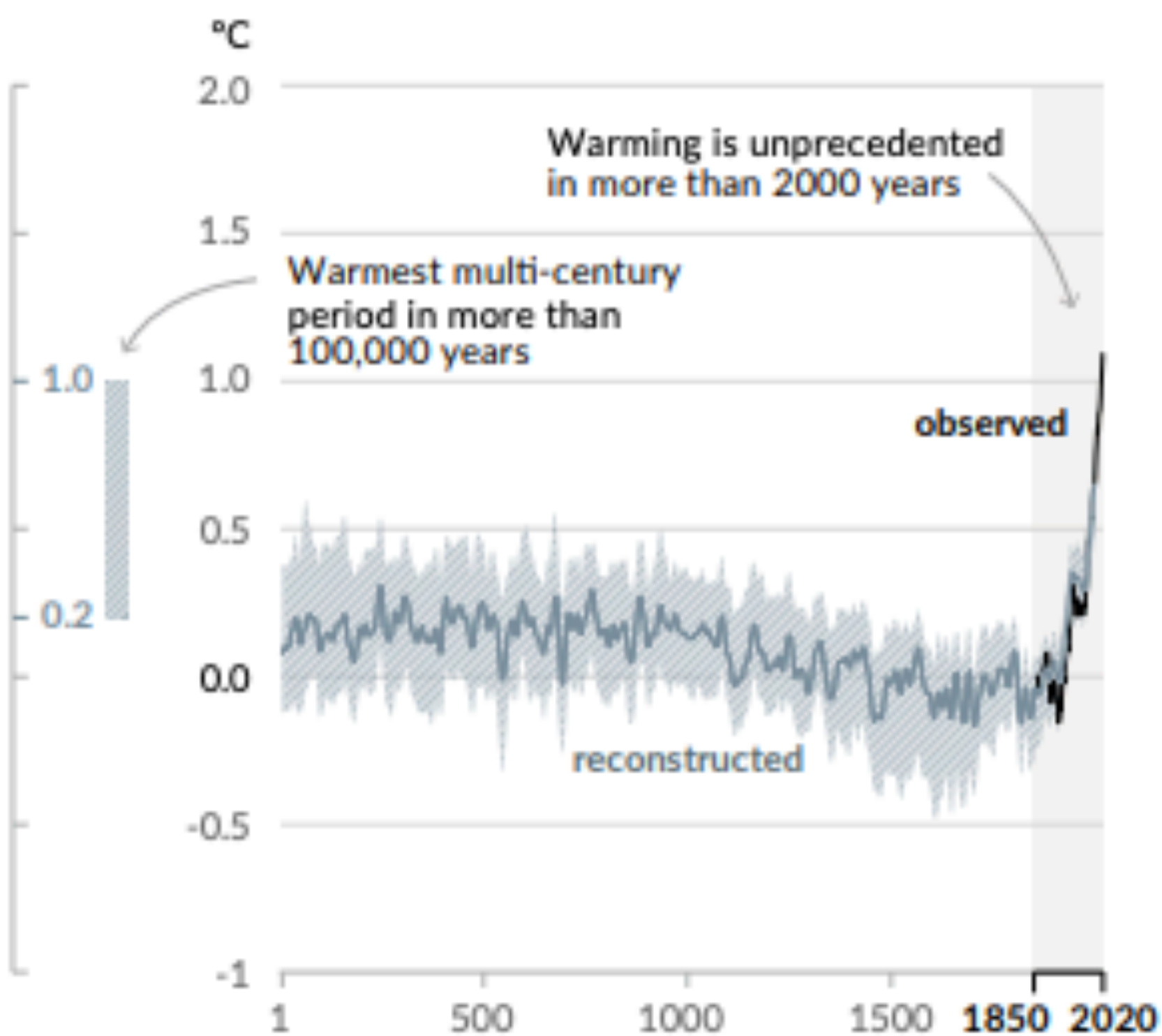
www.gurmanbhatia.com

Human influence has warmed the climate at a rate that is unprecedented in at least the last 2000 years

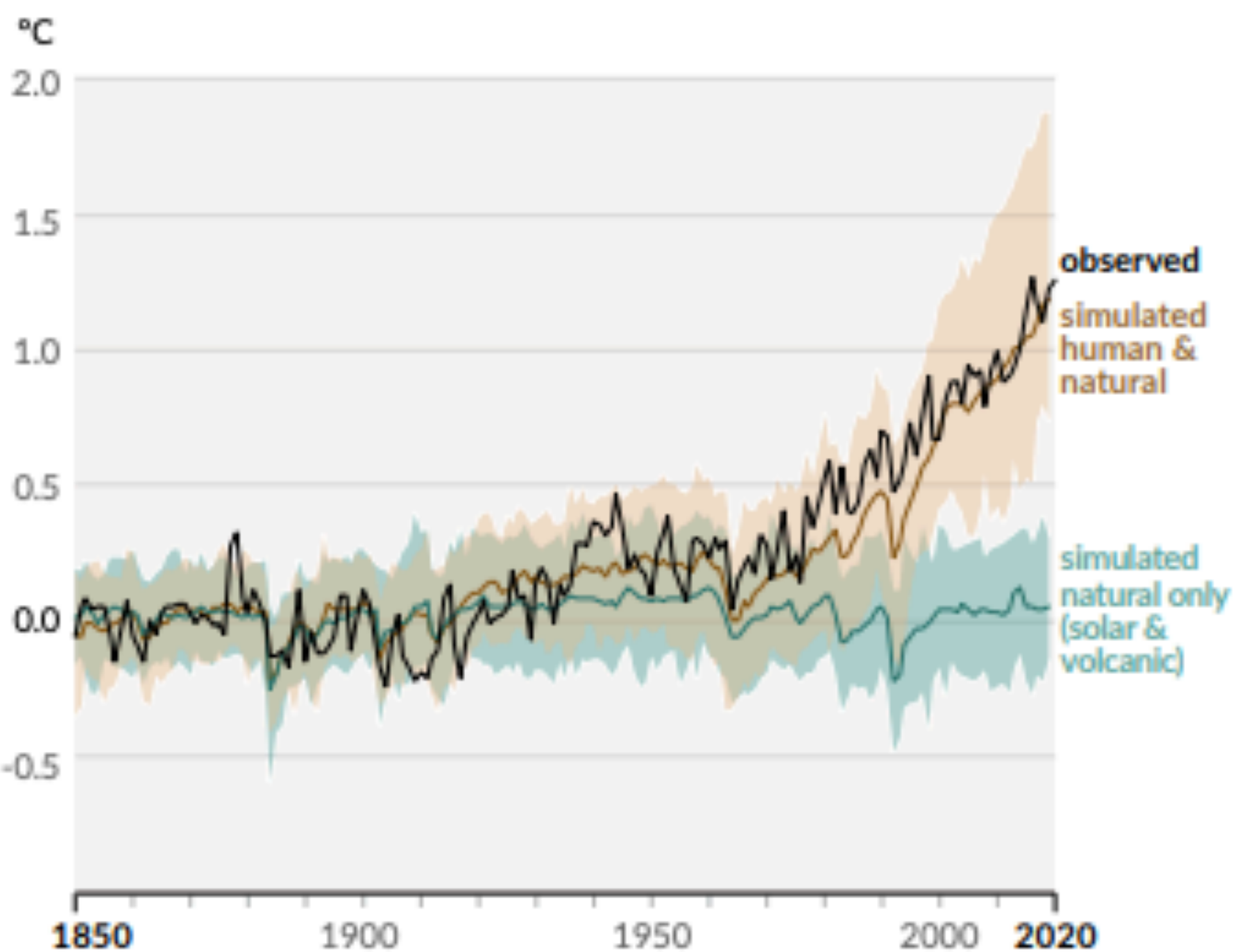
Research

Changes in global surface temperature relative to 1850-1900

a) Change in global surface temperature (decadal average) as **reconstructed** (1-2000) and **observed** (1850-2020)



b) Change in global surface temperature (annual average) as **observed** and simulated using **human & natural** and **only natural** factors (both 1850-2020)

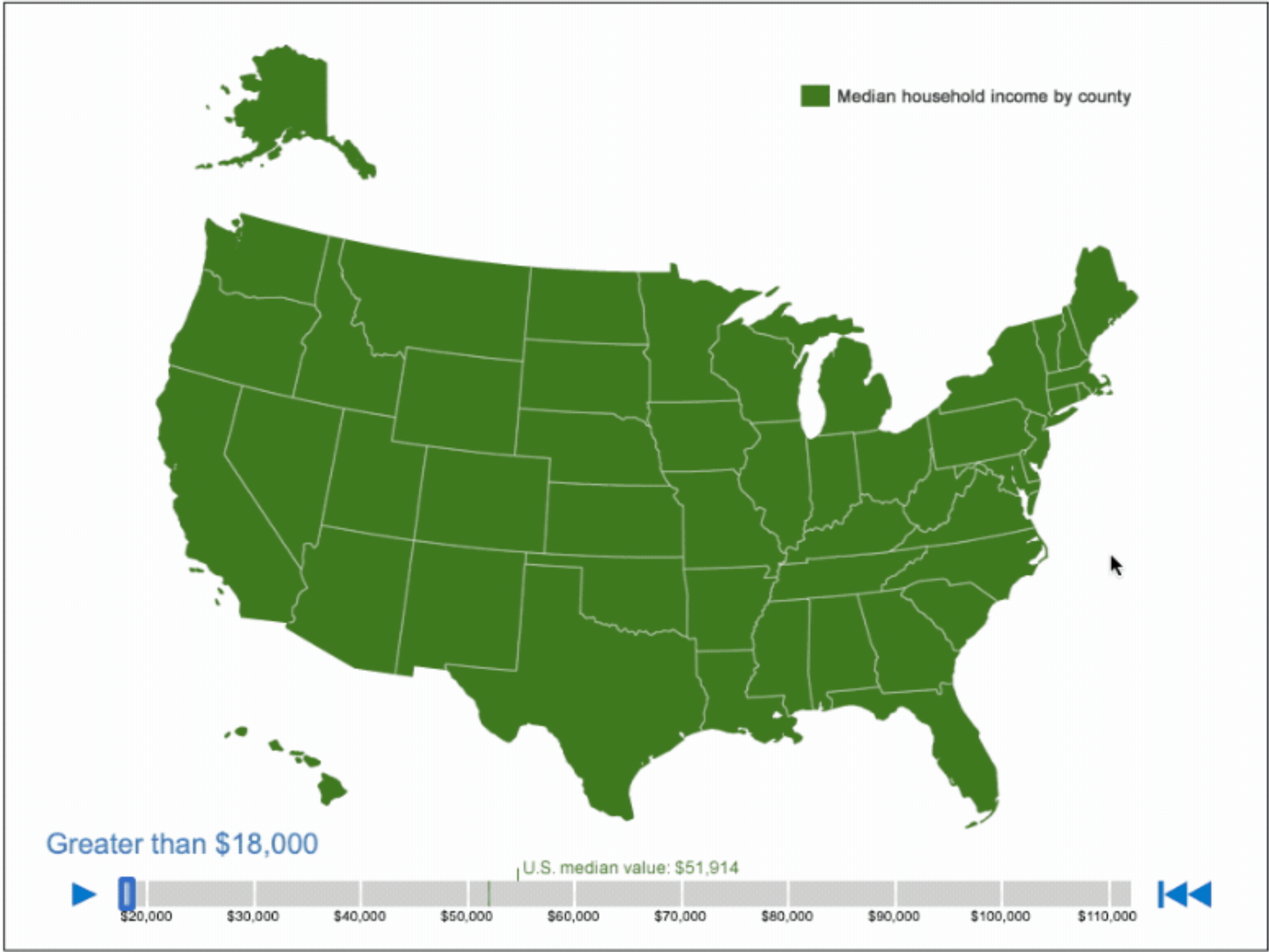


Source: IPCC SPM Report

Islands of High Income

September 20, 2012

Public Policy



Source: The U.S. Census Bureau

Non-Profit

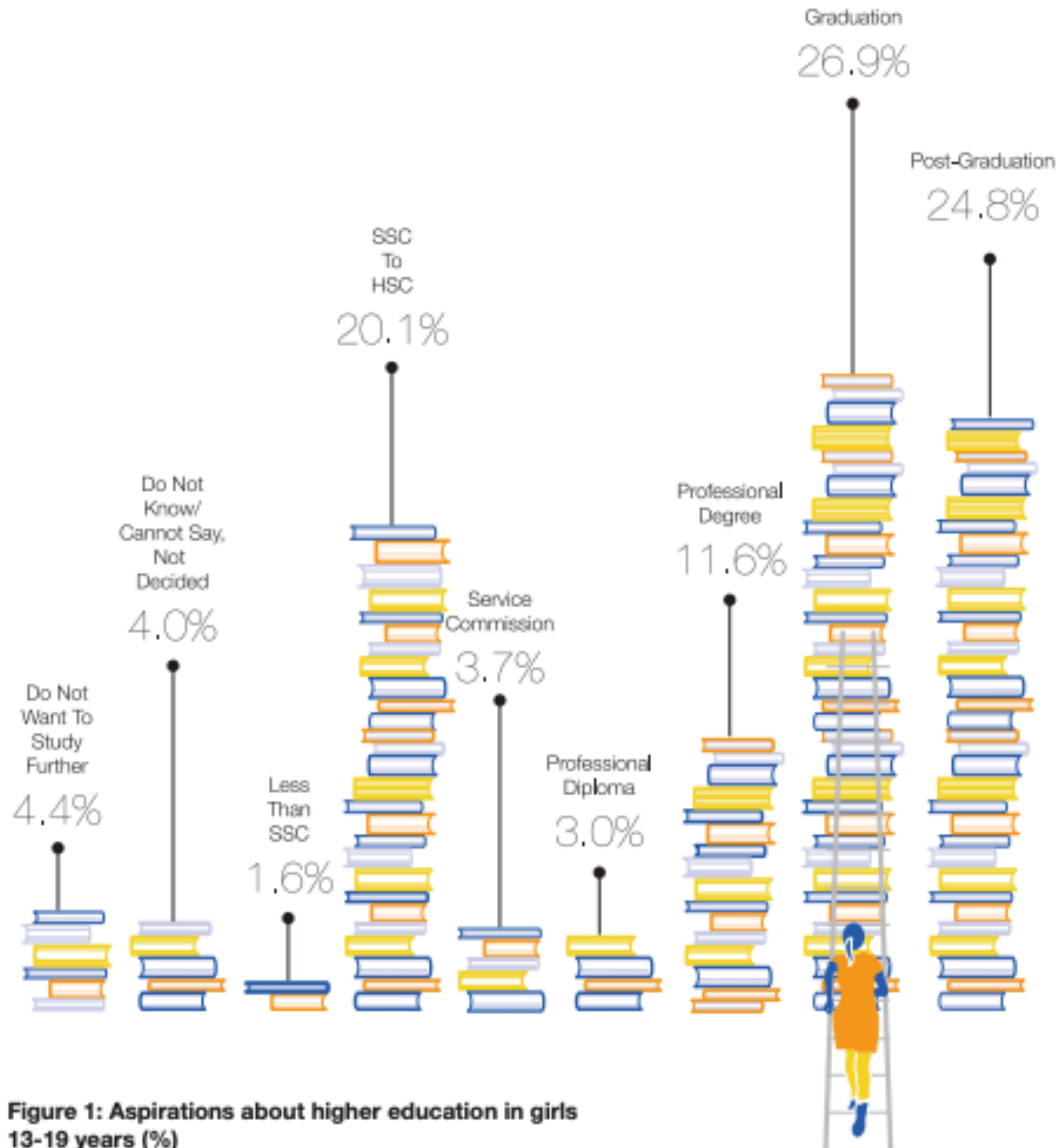
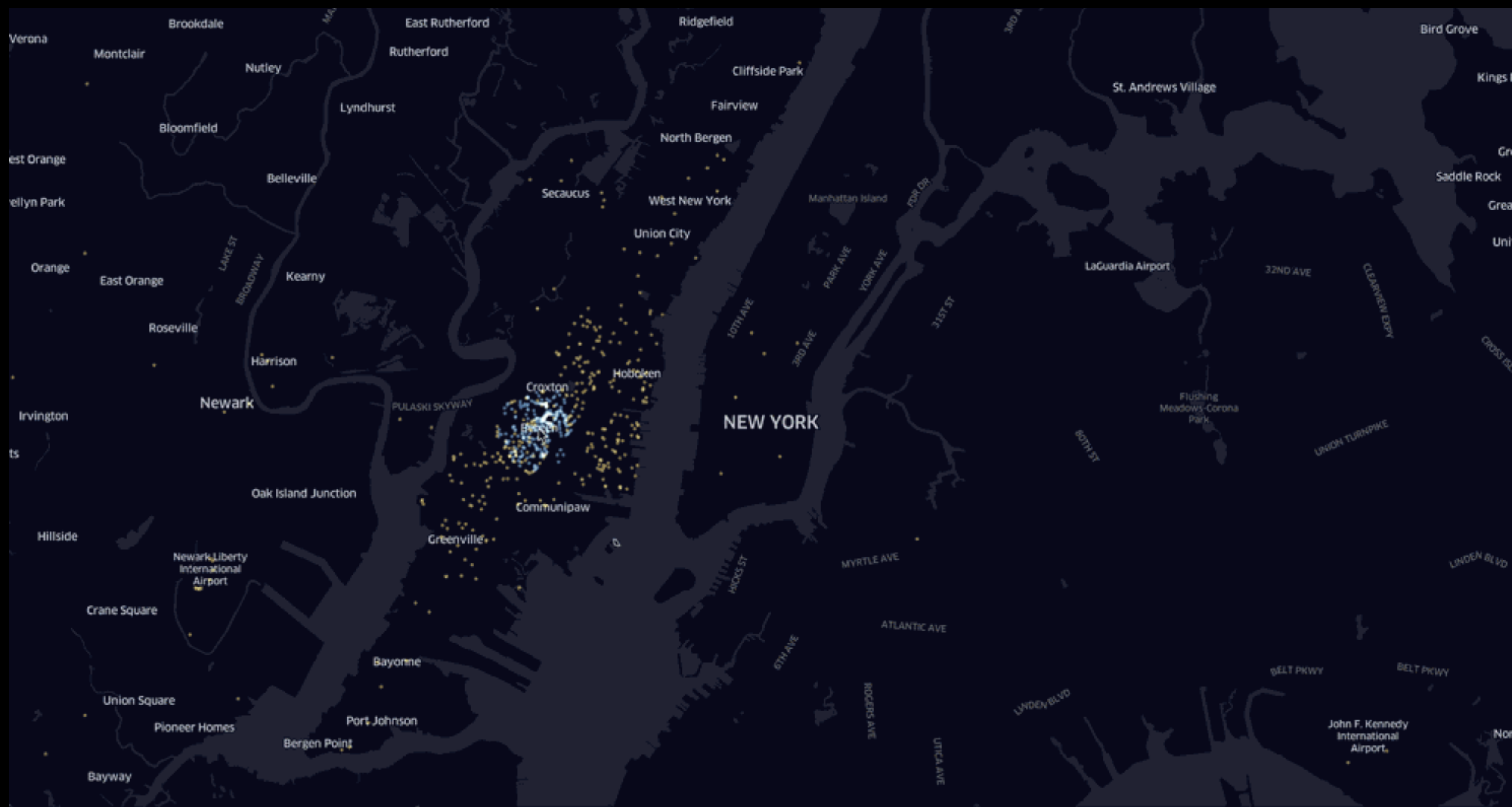


Figure 1: Aspirations about higher education in girls 13-19 years (%)

Source: Teen Age Girls Survey,
Project Nanhi Kali

www.gurmanbhatia.com

Corporate



Dragging the cursor over a given radius area reveals distributions of Uber dropoffs in real-time.

Source: Uber Engineering.

Which means
data visualisation
is catching on...

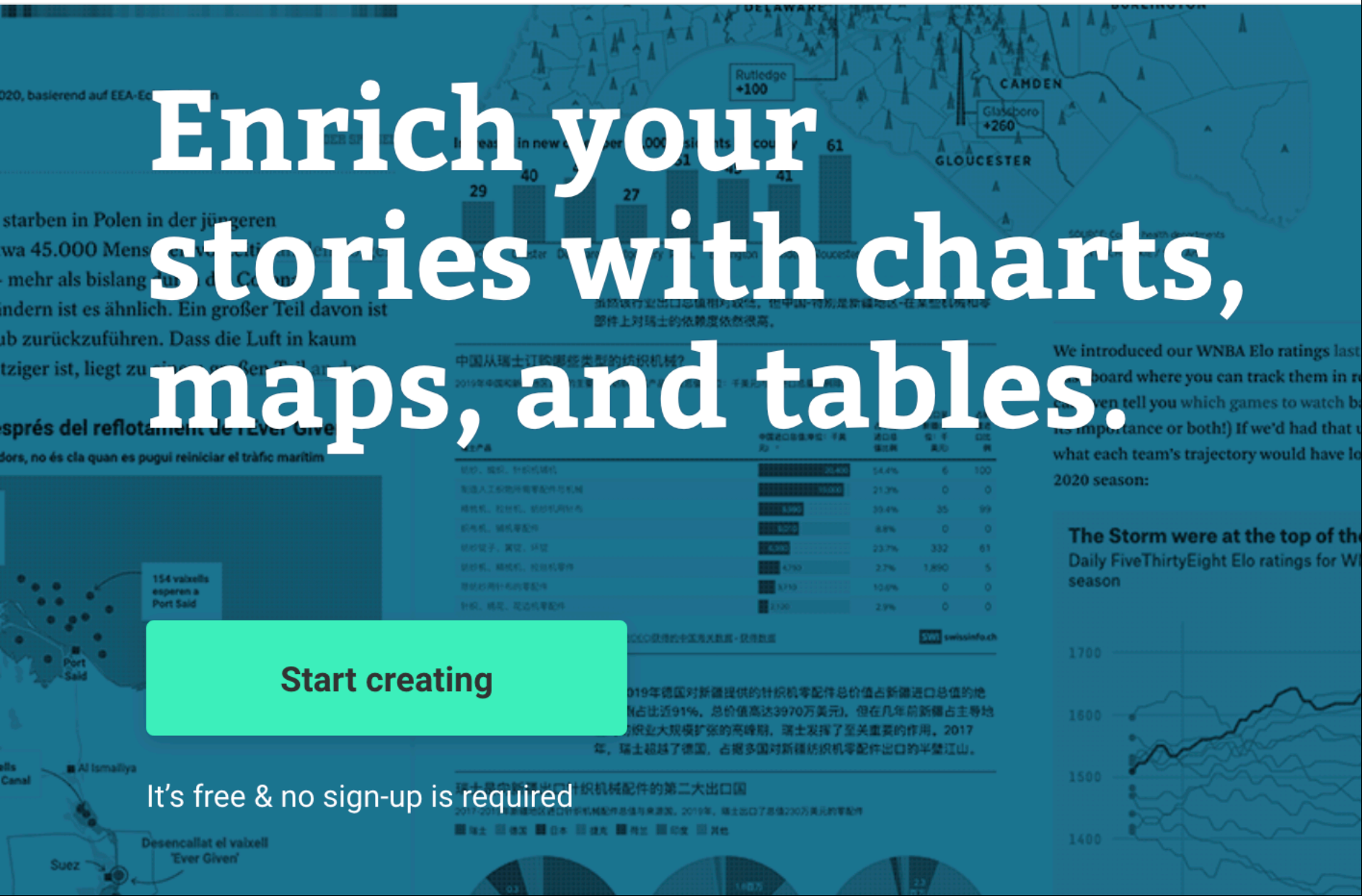


A lot has
happened
recently

1. Better tools

1. Better tools

No-code solutions

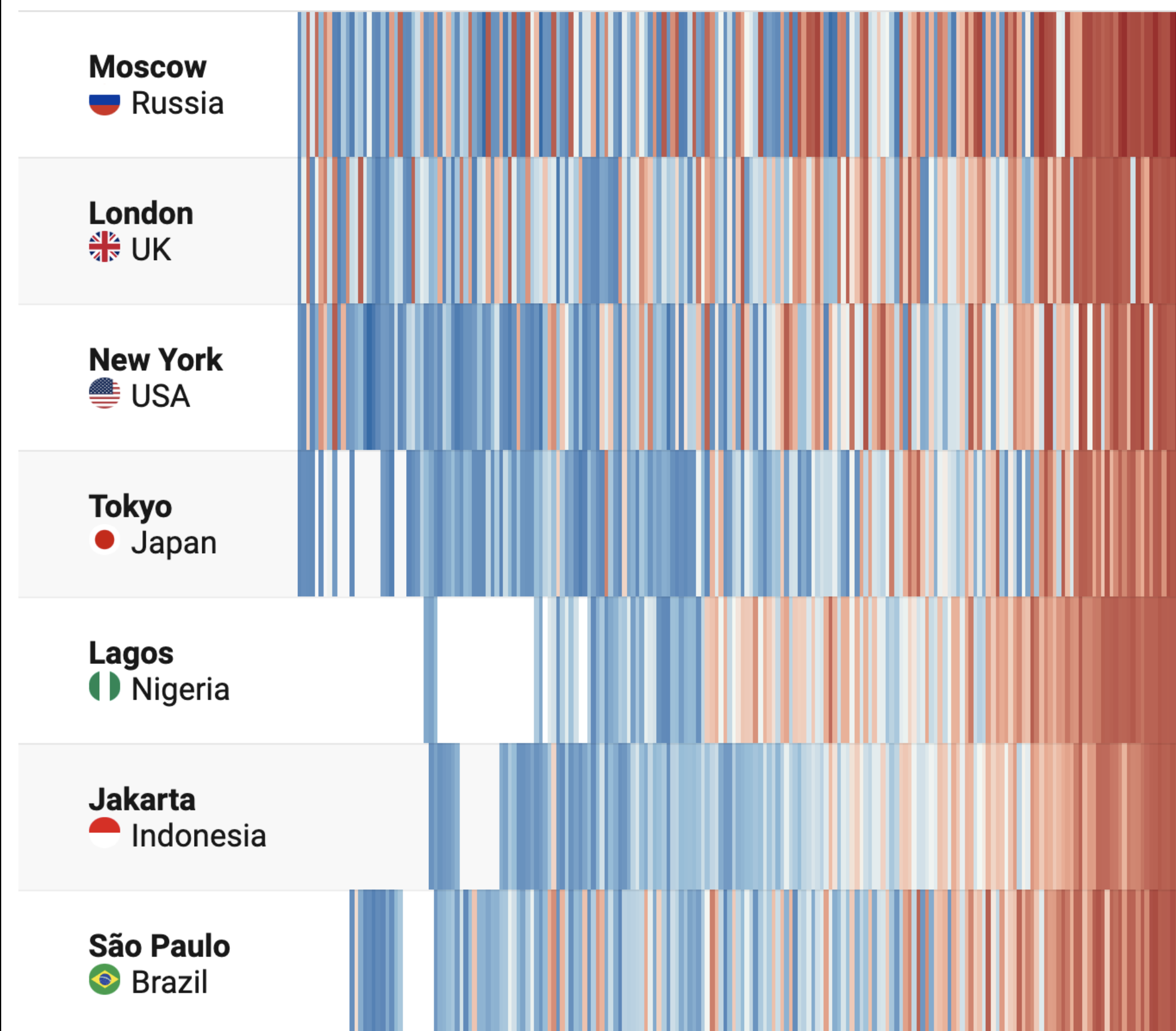


Datawrapper



Flourish

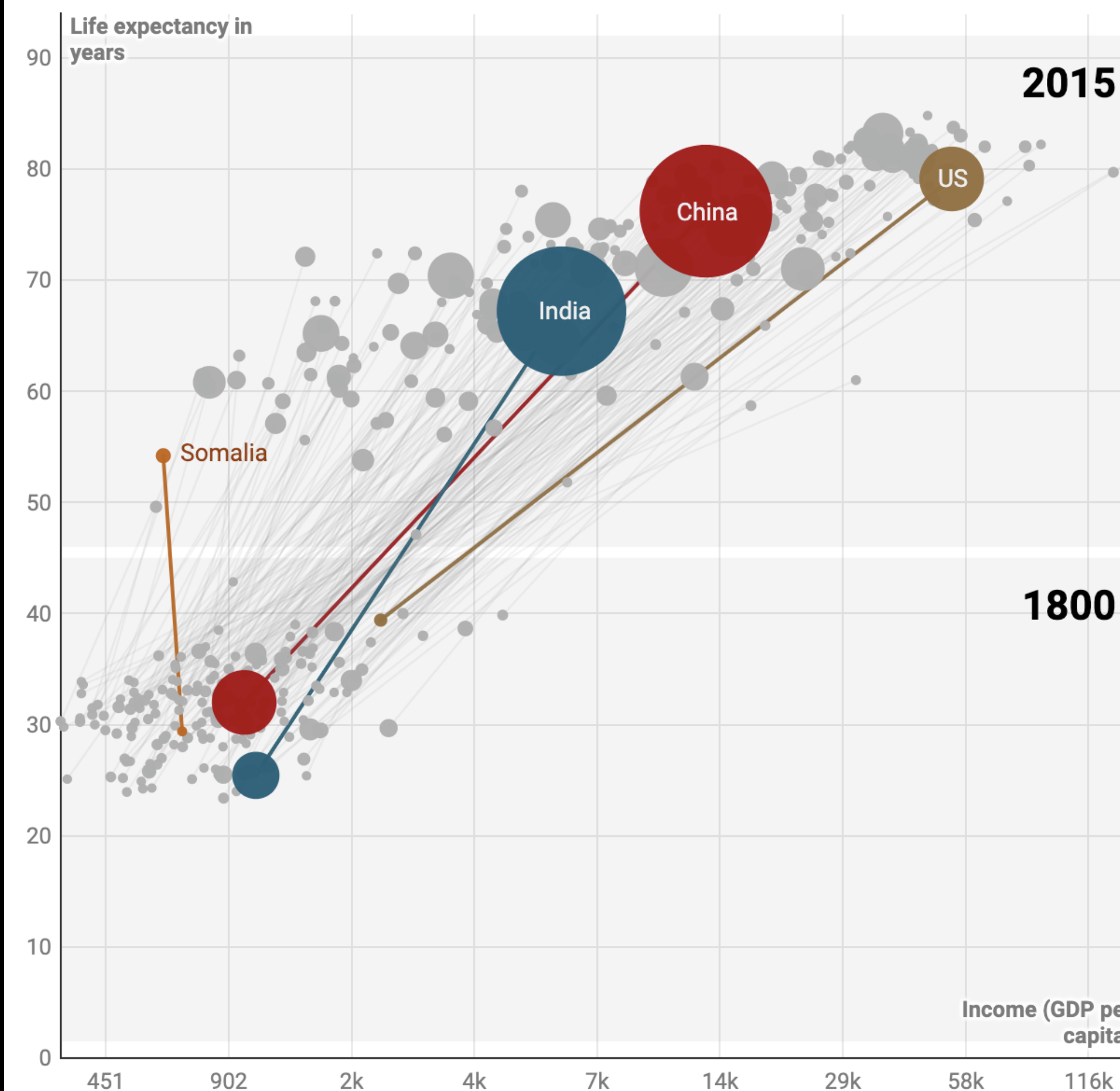
Temperature change around the world (1820–2020)



Temperatures in Celsius, reported as anomalies relative to the Jan 1951–Dec 1980 average.

Every country has a higher life expectancy now than in 1800

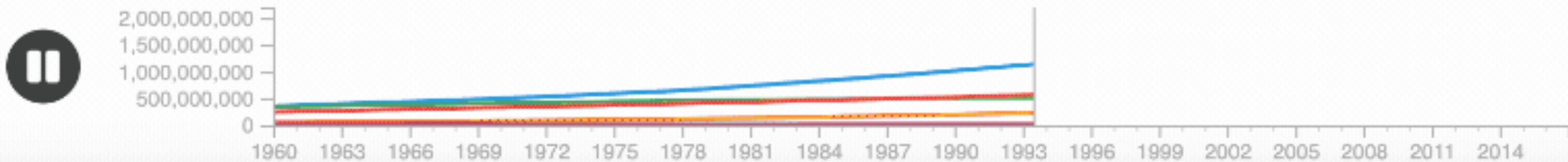
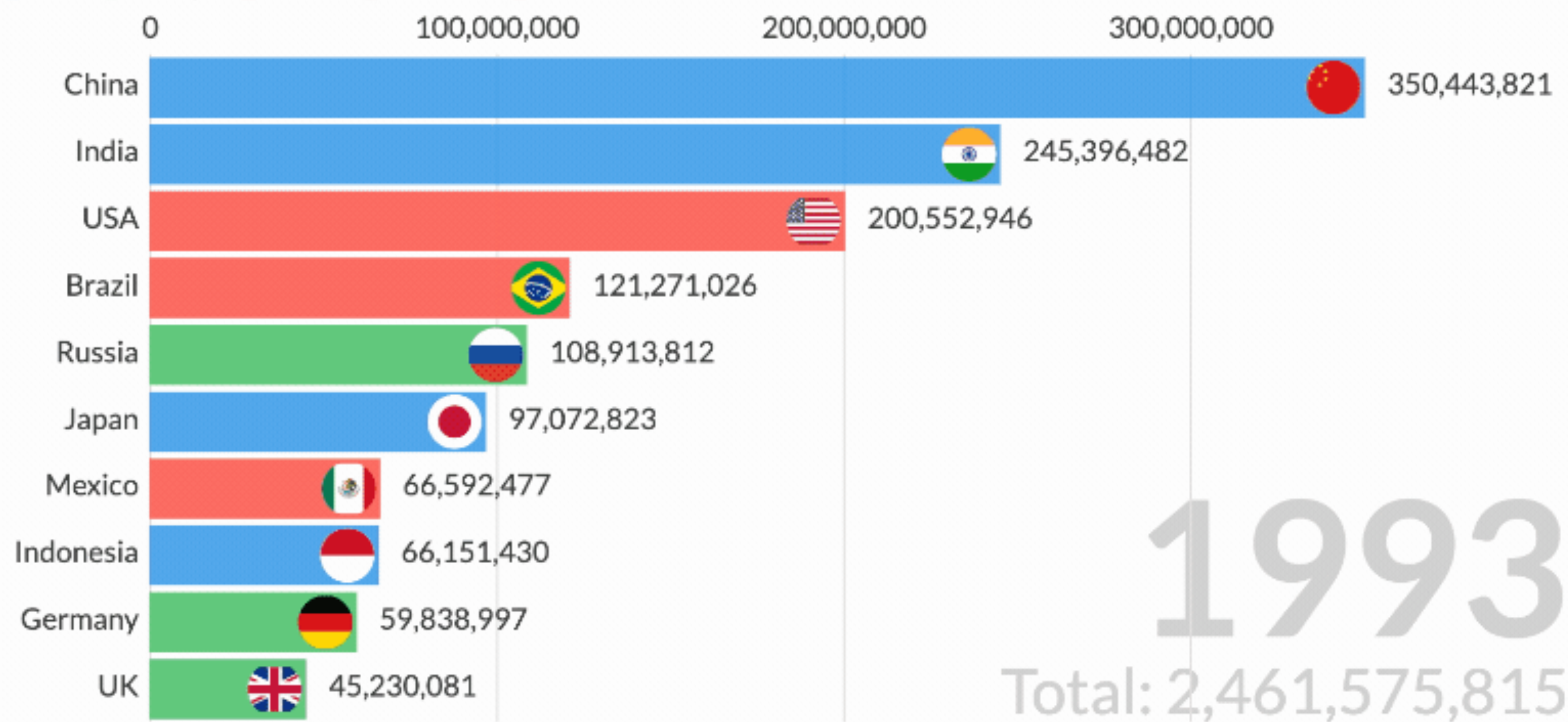
Gross domestic product per person adjusted for differences in purchasing power (in international 2011 dollars) vs life expectancy of newborns, for selected countries in 1800 and 2015. The size of the circle shows the population of the countries.



Urban population by country

Coloured by continent. Click the legend to filter.

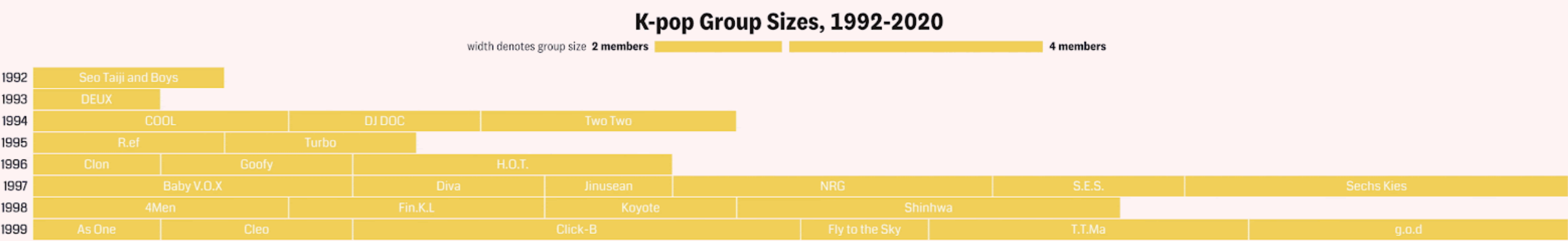
Asia Europe Africa Oceania Americas



members (yes, a group that big exists) might seem alien, or downright excessive. And yet the average size of the top 10 selling K-pop groups of the last decade (like girl group TWICE, pictured above) is 9 members.

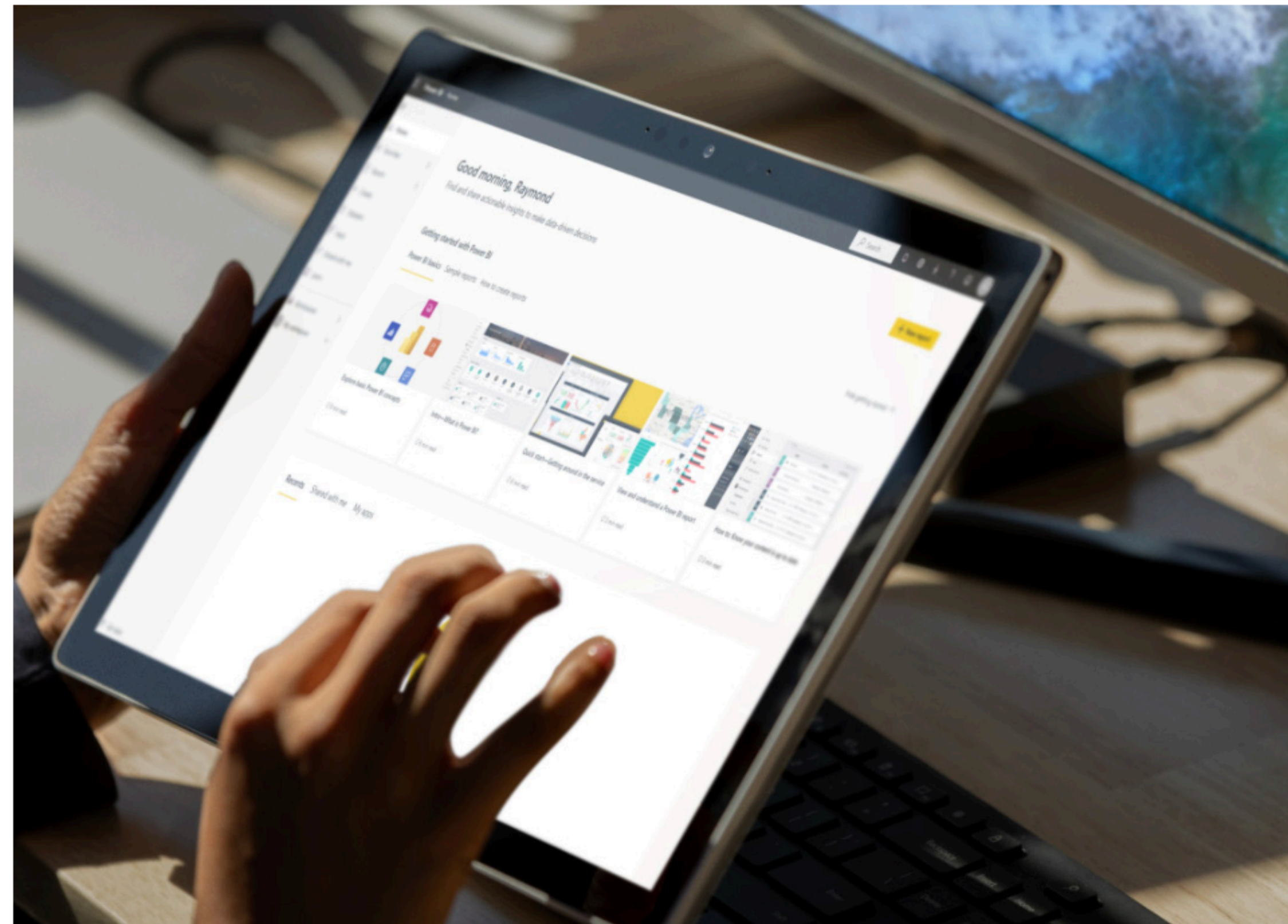
So, how did groups get that large? What about large groups is so appealing? And, what do the sizes of K-pop groups tell us about why K-pop is so popular?

To answer those questions, we tracked trends in group sizes and member roles over modern K-pop’s 30-year history, breaking the numbers down across the industry’s three generations of artists.





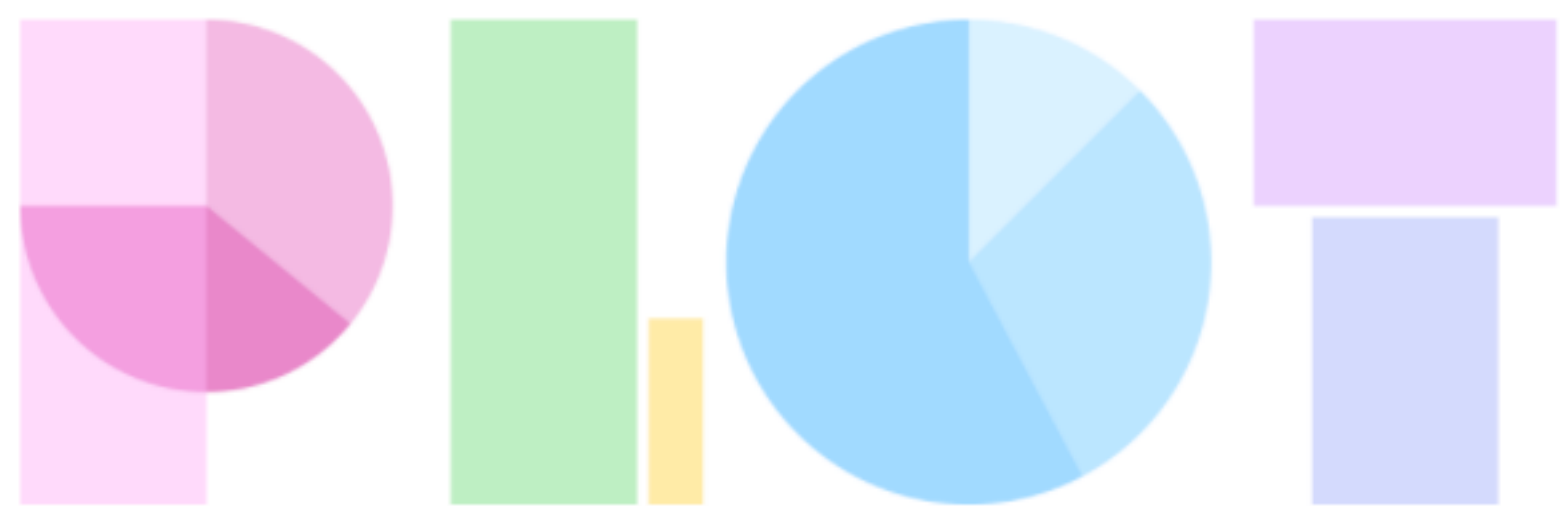
Tableau



PowerBI

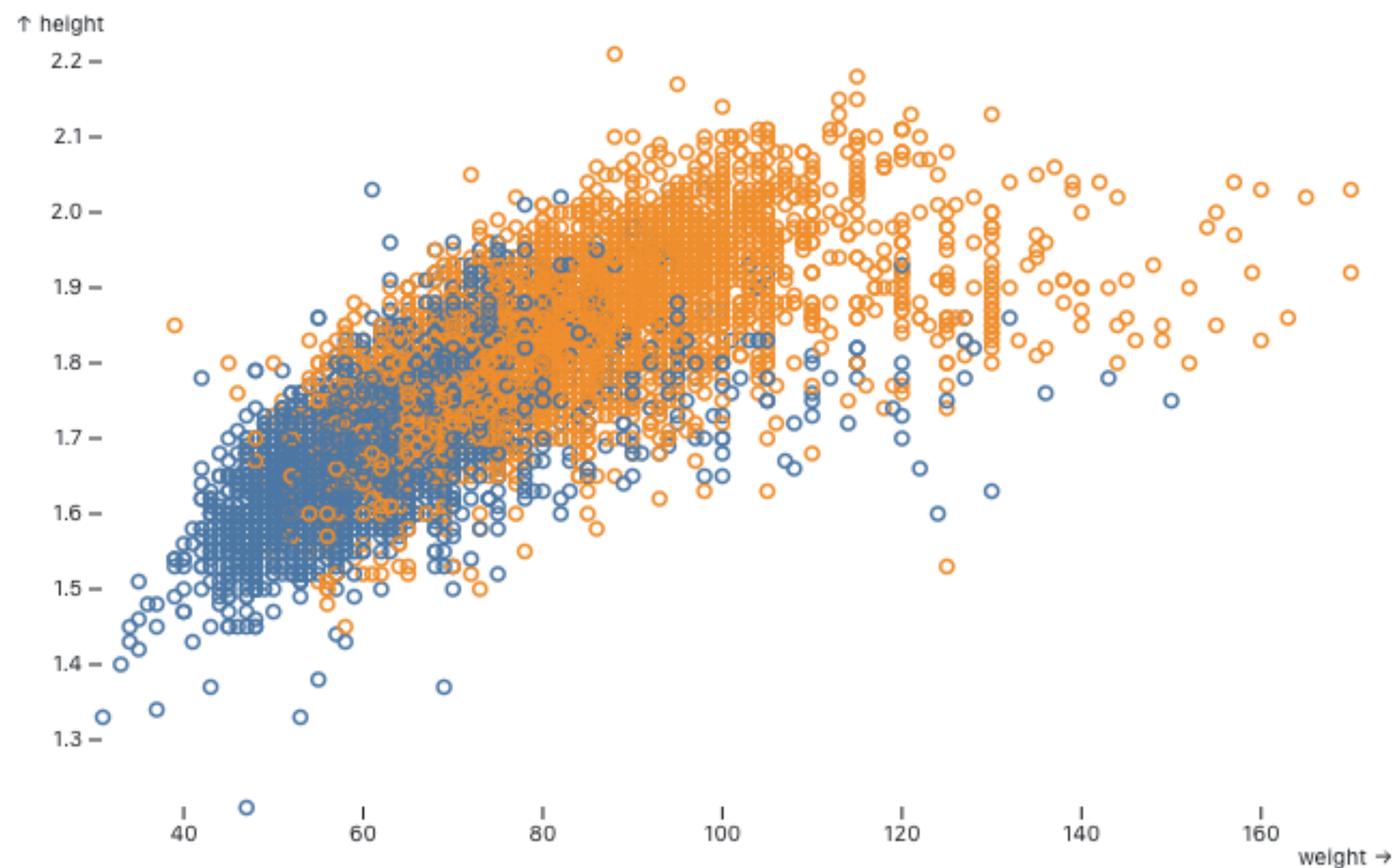
1. Better tools

Low-code solutions



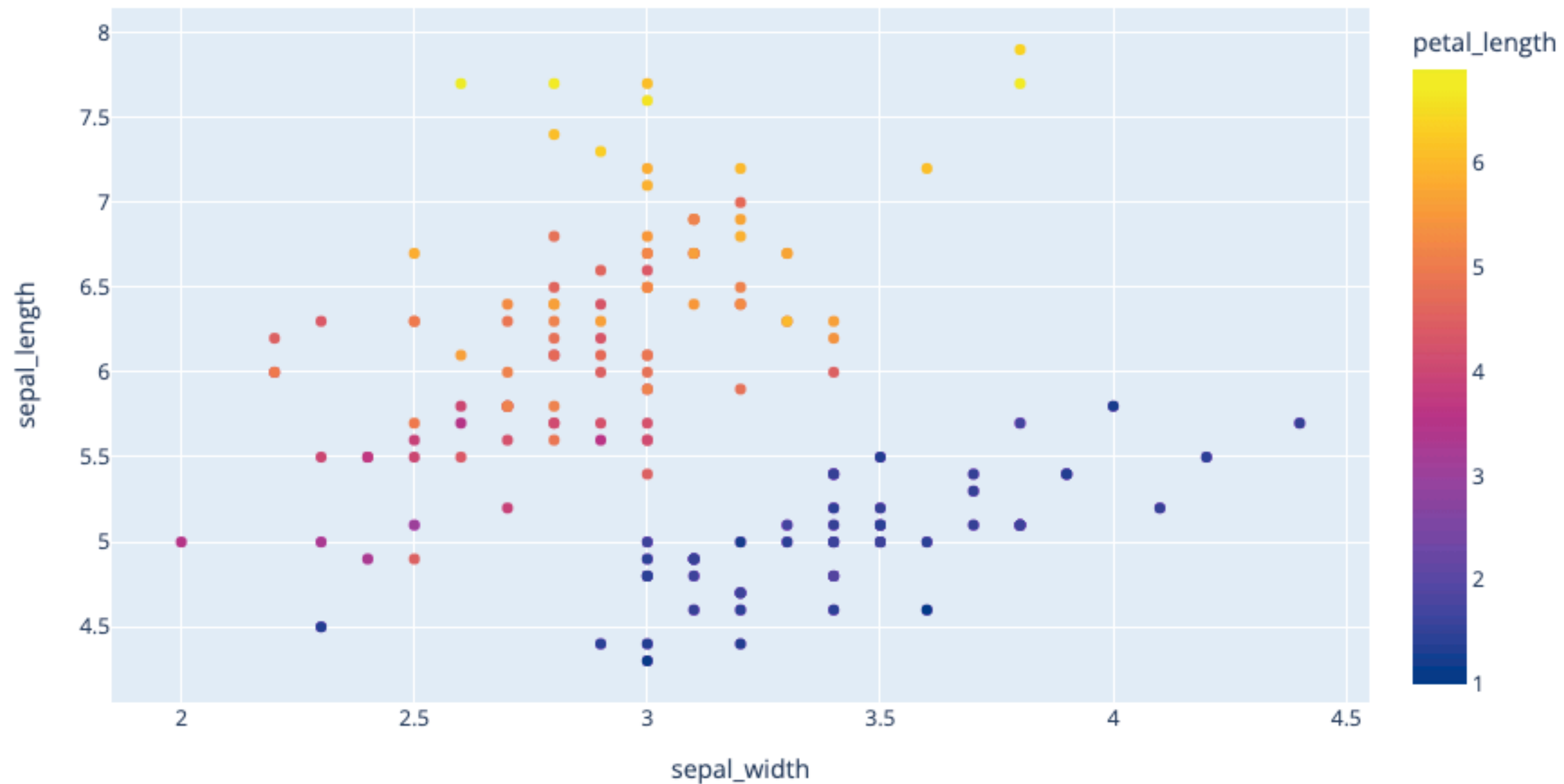
🕒 Observable

female male

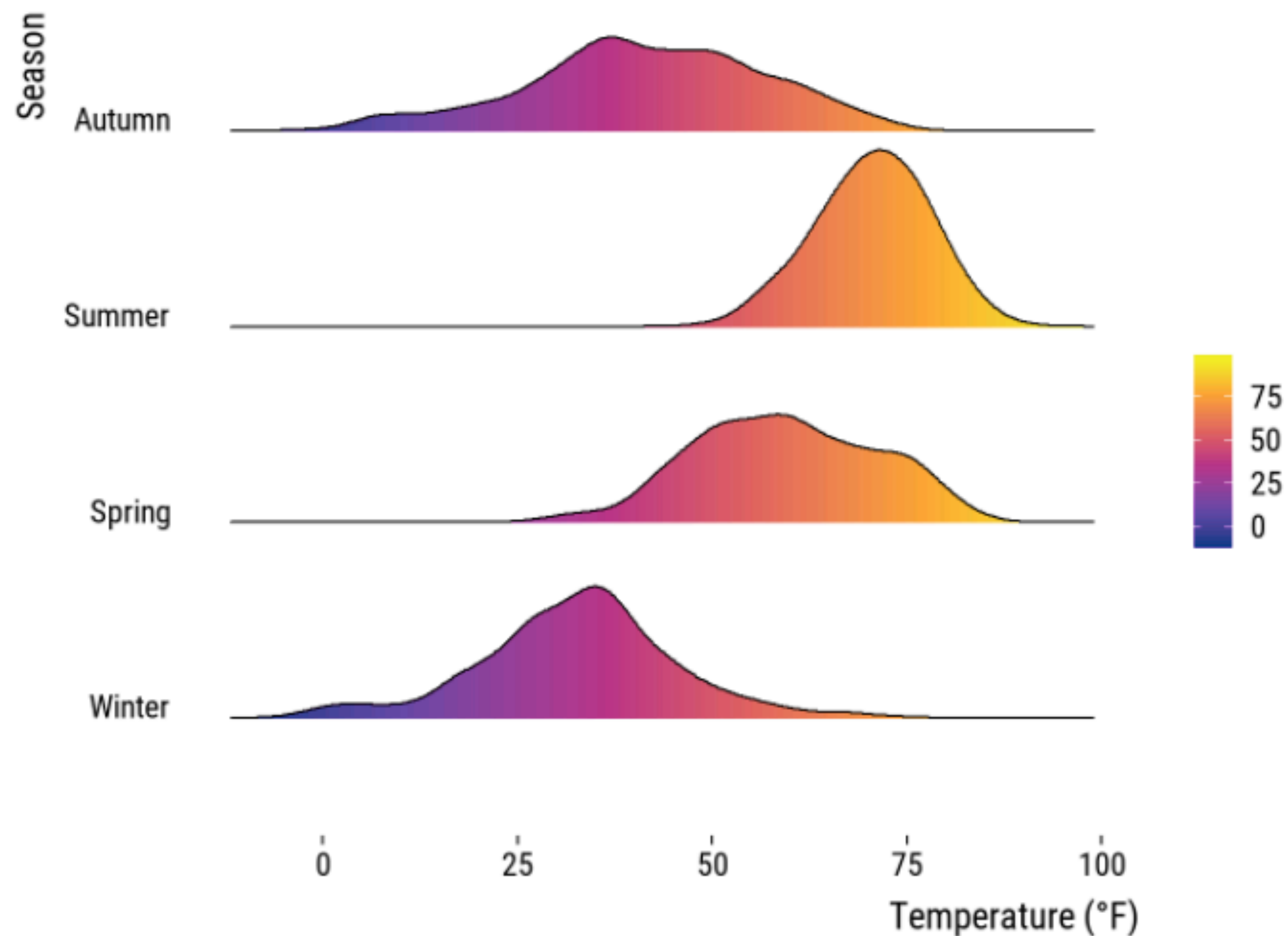


```
dotplot = Plot.dot(athletes, {x: "weight", y: "height", stroke: "sex"}).plot()
```

```
import plotly.express as px
df = px.data.iris()
fig = px.scatter(df, x="sepal_width", y="sepal_length", color='petal_length')
fig.show()
```

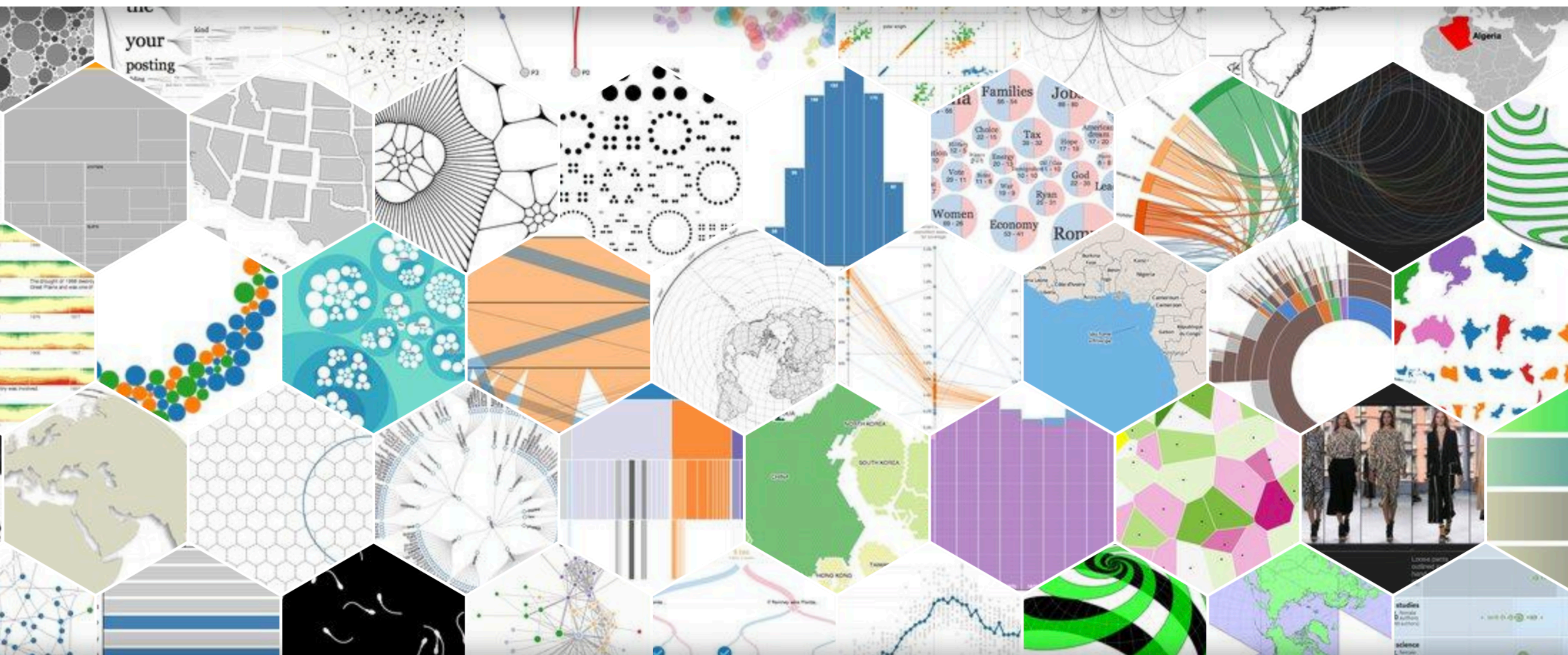


```
ggplot(chic, aes(x = temp, y = season, fill = ..x..)) +  
  geom_density_ridges_gradient(scale = .9, gradient_lwd = .5,  
                              color = "black") +  
  scale_fill_viridis_c(option = "plasma", name = "") +  
  labs(x = "Temperature (°F)", y = "Season") +  
  theme_ridges(font_family = "Roboto Condensed", grid = FALSE)
```





Data-Driven Documents




```

const node = this.selection().node();
const { width: containerWidth } = node.getBoundingClien
const width = containerWidth - margin.left - margin.rig
const height = (p
  this.selection()
    .appendSelect('p.axis-label')
    .text(props.topText)
    .style('padding-left', (props.margin.lef

const xScale = d3
  .rangeRound([0,
  .domain([0, use

const yScale = d3
  .rangeRound([0,
  .domain(data.ma

const plot = this.svg
  .appendSelect('g')
  .classed('plot', true)
  .attr('transform', `translate(${margin.l

const yAxis = plot.appendSelect('g.axis.y'
const xAxisTop = plot.appendSelect('g.axis

xAxisTop.attr('transform', 'translate(0,0)')
  .call(
    d3.axisTop(xScale)
      .ticks(4)
      .tickFormat(d => d * 100)
  );

yAxis.attr('transform', 'translate(0, 0)')
  .call(
    d3.axisLeft(yScale)
      .tickFormat((d) => {
        return props.countryNameGetter(d);
      })
  );
):

```

```

const bars = plot.appendSelect('g.country-container')
  .selectAll('rect')
  .data(data, d => d.country);

```

```

bars
  .enter()
  .append('rect')
  .style('fill', props.rectFill)
  .merge(bars)
  .attr('x', 0)
  .attr('y', d => yScale(d.country))
  .attr('height', d => yScale.bandwidth())
  .attr('width', d => xScale(d.perPop));

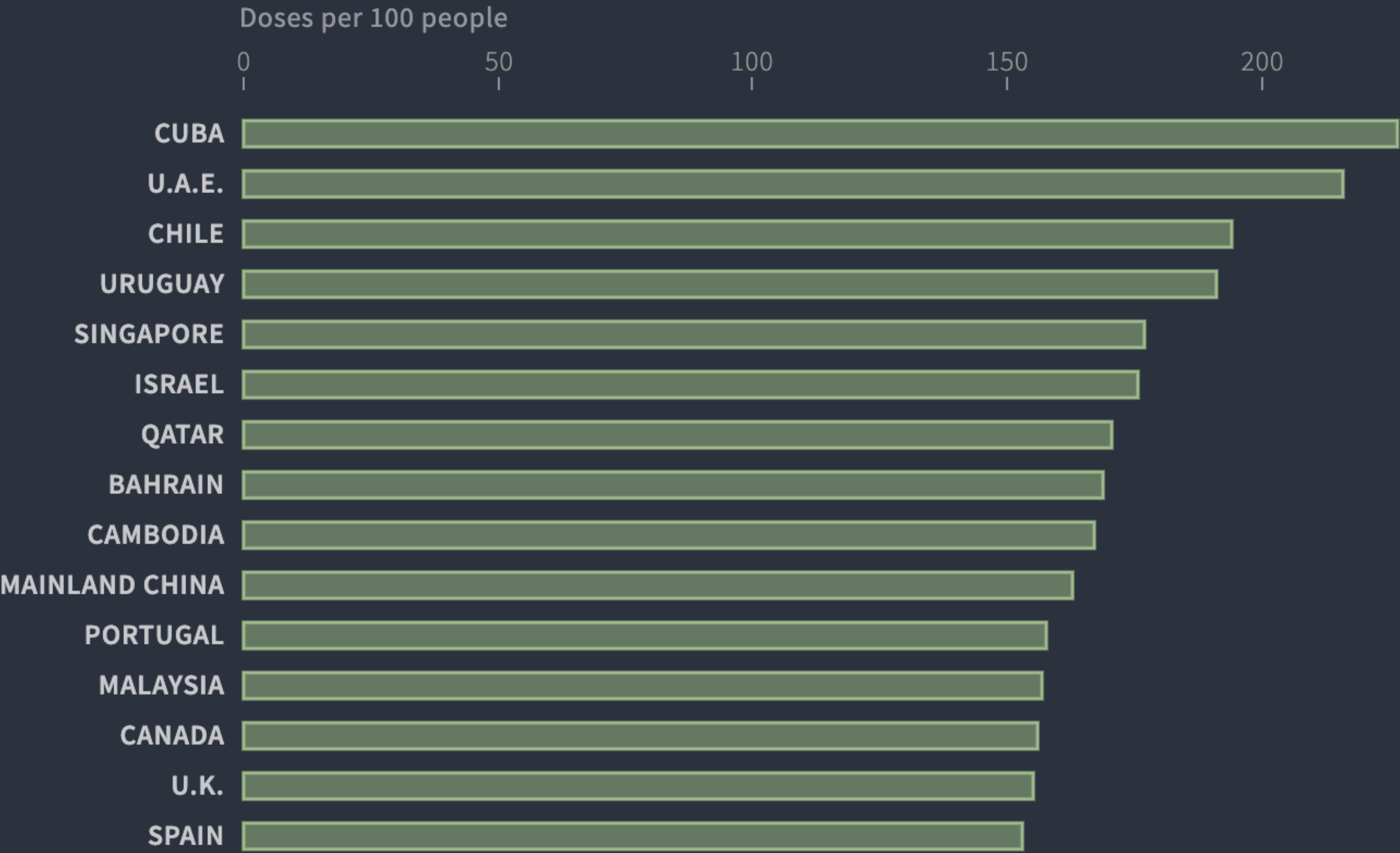
```

```

bars.exit()
  .remove();

```

Countries reporting the most doses administered per population*



*Countries with more than
1 million people

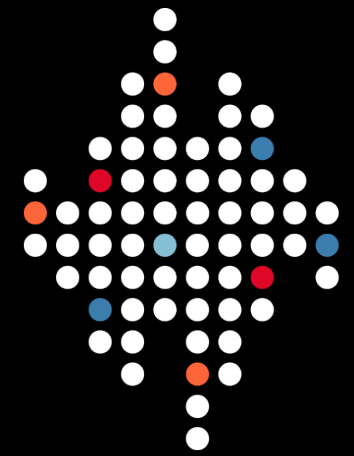
2. A stronger community



Data Visualization Society
datavizsociety.slack.com



16,211



MakeoverMonday



Topi Tjukanov
@tjukanov

🎉 Categories for [#30DayMapChallenge](#) 2021 🎉

Join in to create maps around these themes. Use your creativity and share your results with others! Starts 01/11/2021.

Workout Wednesday

A weekly challenge to re-create data-driven visualizations.



Viz for
Social Good

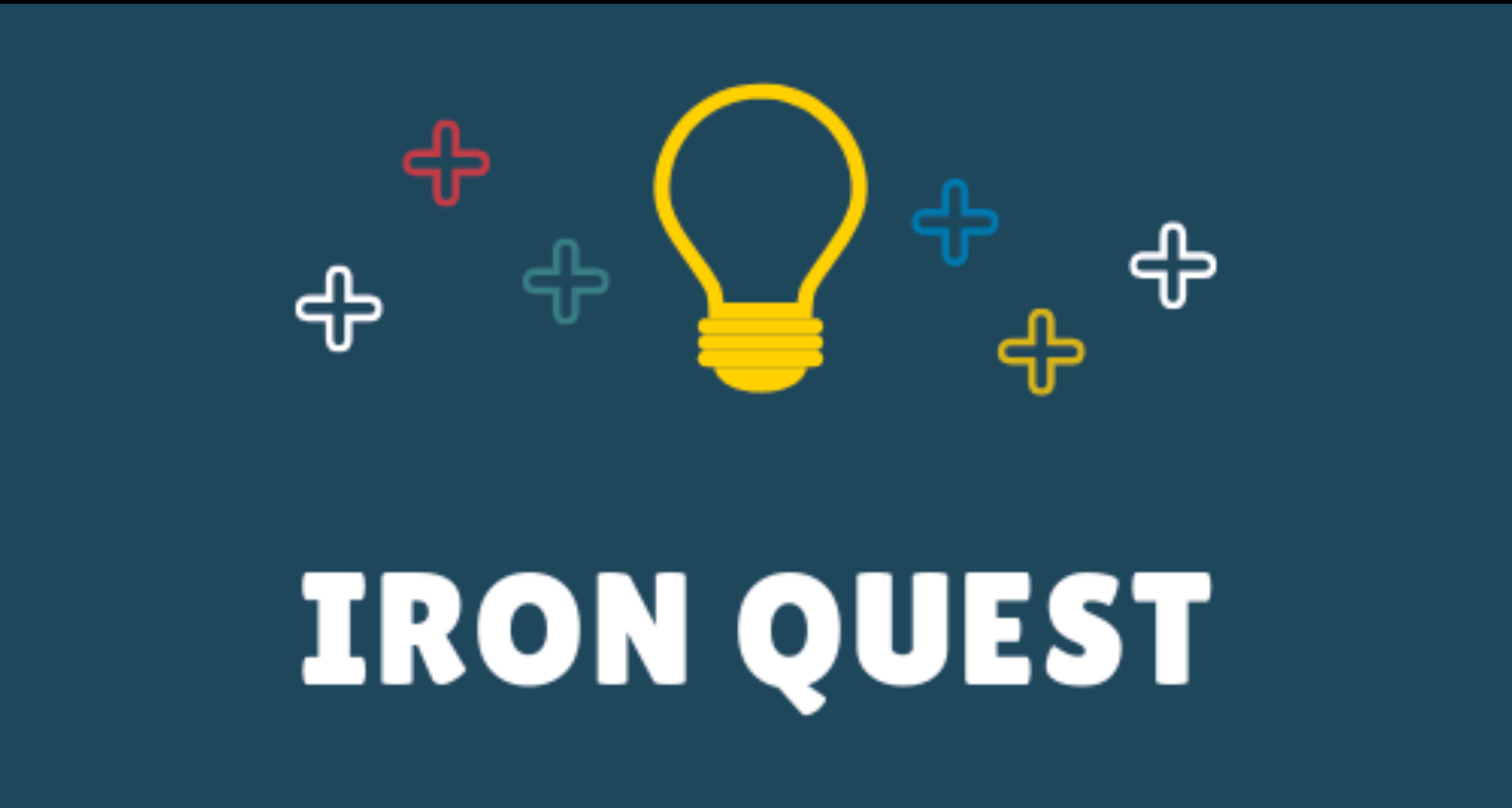
TidyTuesday

A weekly data project in R from the R4DS online learning community


variables

observations

values


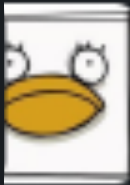


Real World Fake Data



d3

d3js.slack.com



8,112

3. A wealth of cheap/free online resources to learn

Welcome to the Datawrapper Academy!

Here you can learn how to build charts, maps & tables with Datawrapper.

Search



Welcome!

Your first chart, what's Datawrapper & a tour through it.

[10 articles](#)



Bar charts

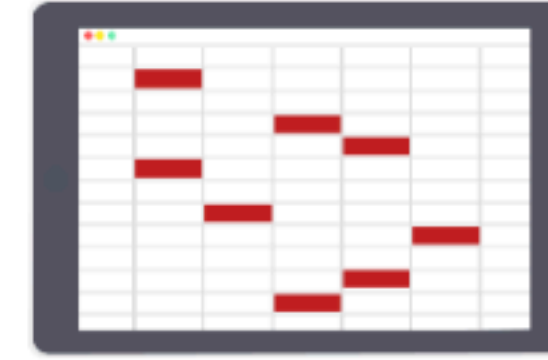
Bar charts, stacked, bullet, split & grouped bar charts.

[13 articles](#)



Flourish basics

26 articles



Editing data

9 articles



Flourish stories

8 articles



Line & Area charts

Line charts, area charts and slope charts

[15 articles](#)



Pie & Donut charts

Pie charts, donut charts, multiple pie charts & multiple donut charts

[12 articles](#)



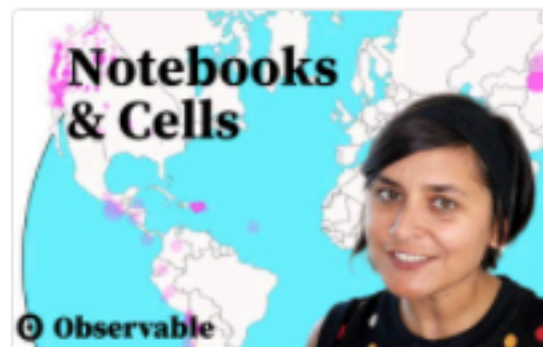
Scatter plots

Create and customize scatterplots, set annotations & add custom lines

[11 articles](#)

Video series: Learning Observable

Understand Observable core concepts & learn basic techniques for working with notebooks through these explanatory videos.



1. Notebooks & Cells



2. Reactive Markdown & HTML



3. Reactive Dataflow



4. User Input & Interactivity



5. JavaScript and Observable



6. Imports & Remixing Cells



7. Getting Data into a Notebook



8. Sharing, Publishing & Embedding



Shirley Wu
@sxywu

- Profile
- Notebooks
- Collections

Workshop: Introduction to D3.js

Learn how to work with SVG and foundational D3 concepts by creating a fun data visualization! To access the full workshop, including video lessons, join Frontend Masters: <https://frontendmasters.com/courses/d3/>

Showing all 9 listings

1.

Introduction to SVG and D3.js
Shirley Wu
Aug 11, 2020 · ❤ 238 💬 4
2.

1. Draw a flower petal on the screen
Shirley Wu
Aug 11, 2020 · ❤ 71 💬 7
3.

2. Select existing petal(s) and bind movie data
Shirley Wu
Aug 11, 2020 · ❤ 20 💬 13
4.

3. Create a petal for each movie
Shirley Wu
Aug 11, 2020 · ❤ 7 💬 4

FULLSTACK D3

and DATA VISUALIZATION

The *Fullstack D3* course is the complete guide to D3. With dozens of code examples showing each step, you will gain new insights into your data by creating visualizations in this self-paced online course.

Watch the First Day Free

Enroll Now →

A Quick How-to on Labelling Bar Graphs in ggplot2

Bar charts are likely the most common chart type out there and come in several varieties. Most notably, direct labels can increase accessibility of a bar graph. I got a request how one can add percentage labels inside the bars and how to highlight specific bars with {ggplot2}. This short tutorial shows you multiple ways how to do so.

Posted by Cédric • Monday, July 5, 2021

Visualizing Distributions with Raincloud Plots (and How to Create Them with ggplot2)

Raincloud plots, that provide an overview of the raw data, its distribution, and important statistical properties, are a good alternative to classical boxplots. In this tutorial, I highlight the potential problem of boxplots, illustrate why raincloud plots are great, and show numerous ways how to create such hybrid charts in R with {ggplot2}.

Posted by Cédric • Sunday, June 6, 2021



Always coding. Passionate about design. Worried about nature. Proud dad.



QUICK LINKS

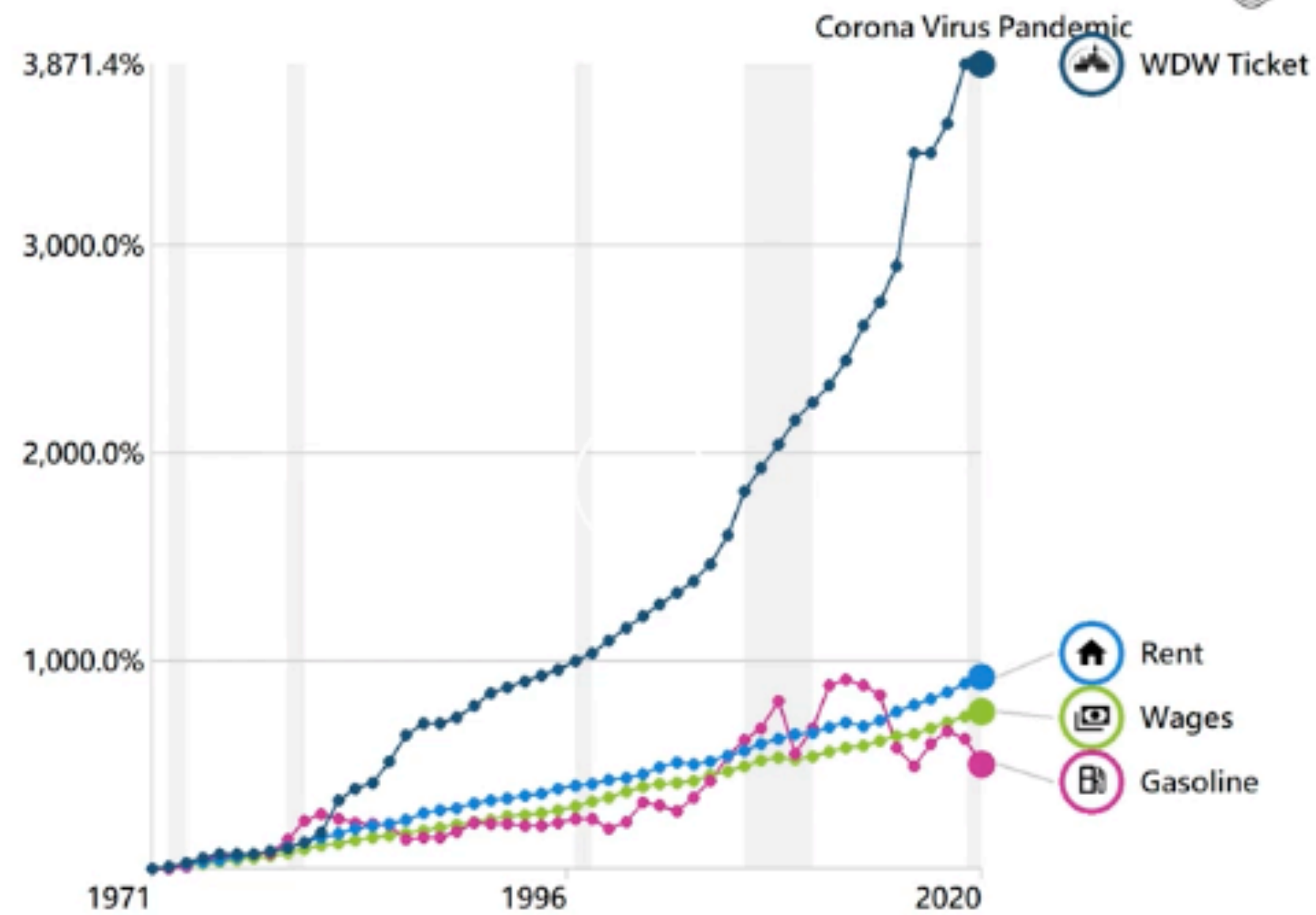
- ggplot2 Tutorial
- Evolution of a ggplot

4. More visual platforms



Posted by u/PieChartPirate OC: 59 16 days ago 🏠 2 🗨️ 18 🔄 13 💰 18 & 8 More
57.2k OC [OC] Walt Disney World Ticket Price Increase vs Wages, Rent, and Gasoline

Walt Disney World Price Increase vs Wages, Rent, and Gasoline



2020



About Community

DataIsBeautiful is for visualizations that effectively convey information. Aesthetics are an important part of information visualization, but pretty pictures are not the sole aim of this subreddit.

16.4m
Members

4.5k
Online

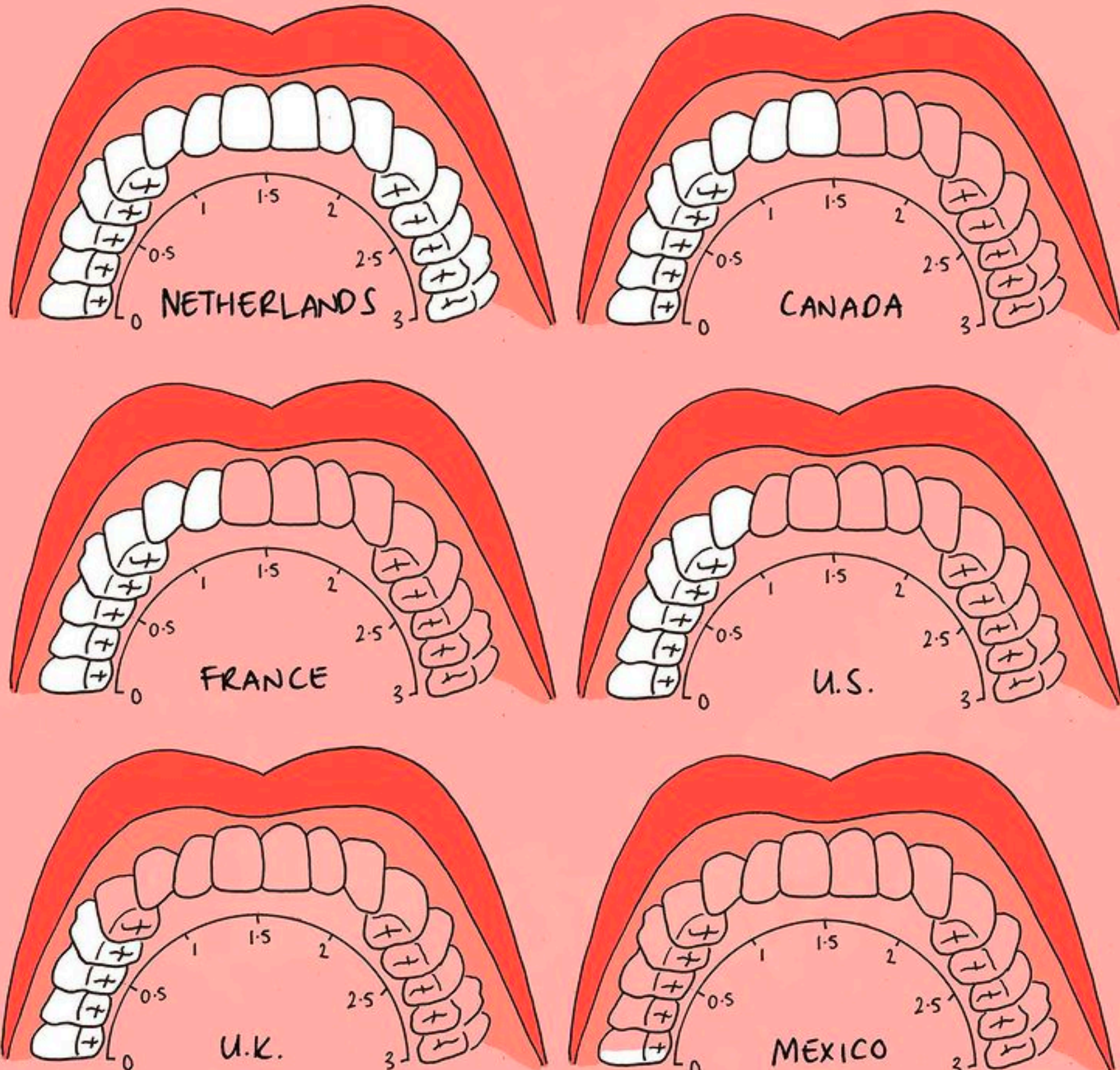
Created Feb 14, 2012

r/dataisbeautiful topics

Science

3.2k Comments Share Save ...

Average number of dentist consultations per person per year



monachalabi In Brooklyn one summer, I was told two very different things by two very different dentists. Which one could I trust? I explore a toothy little mystery in episode 3 of the podcast Am I Normal.

Source: OECD data 2018-19

6 d



monachalabi ID: There's an outline of red lips and empty teeth against a pink background, the more teeth that are colored in white, the more that people in that country go see a dentist. The data shows: Netherlands 3 dentist consultations on average per year, Canada 1.5, France 1.3, United States 1.1, United Kingdom 0.7 and Mexico 0.1

6 d 171 likes Reply



flotsam @researchremora · Oct 23

A population density map of long, slinky Chile. Took a bit of time to find an angle I like.

#rayshader adventures, an #rstats tale



19

430

2.1K

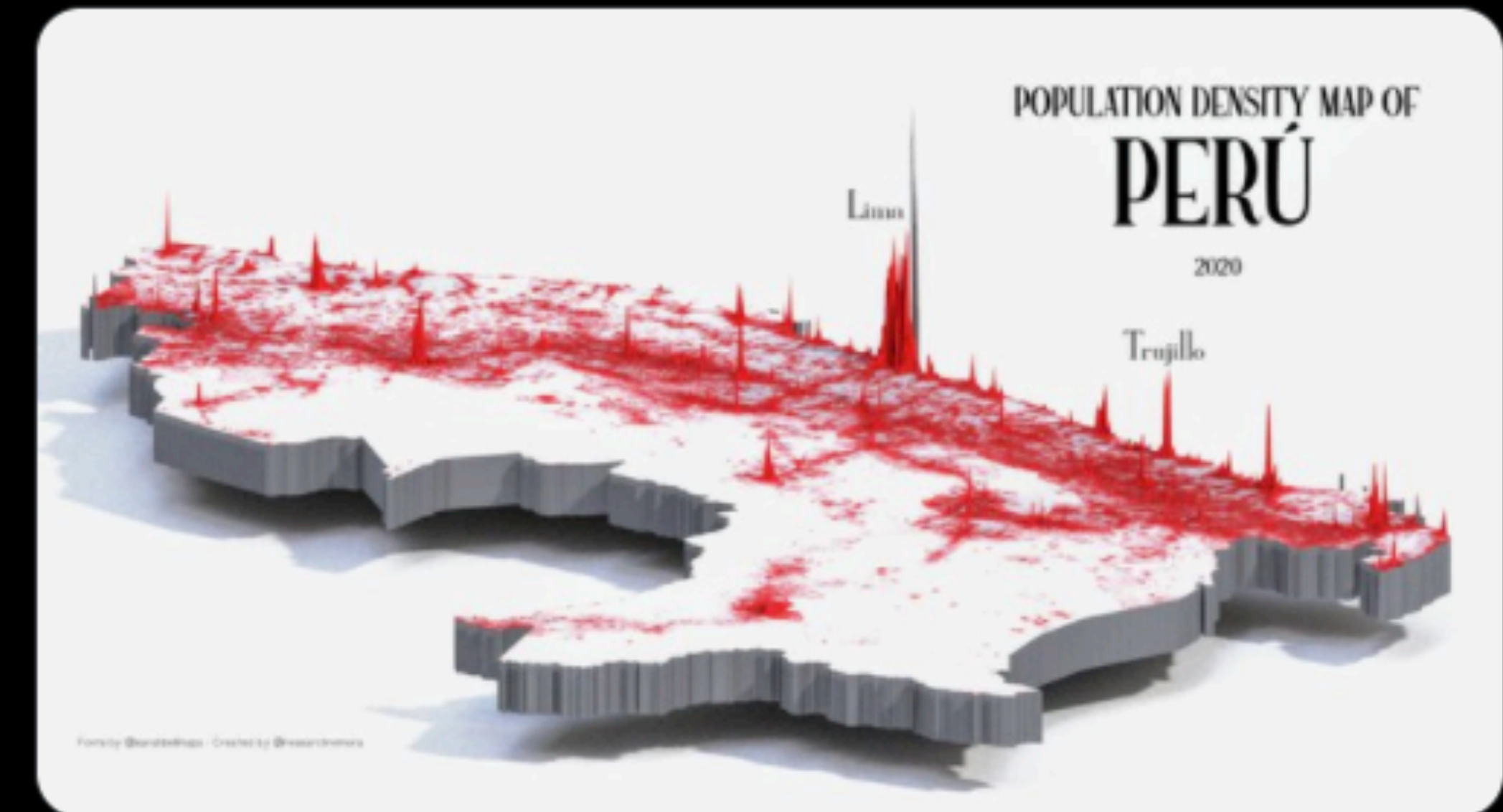


flotsam @researchremora · Oct 21

Today's population density map is Perú.

It was supposed to be México but there's something strange with the data —Monterrey has a huge spike, way higher than Ciudad de México, which seems wrong... (see next tweet)

#rayshader adventures, an #rstats tale



3

21

83





Craig Taylor • Following
Senior Data Visualisation Design Manager
5d •

City Dome | Tokyo, Shinjuku District

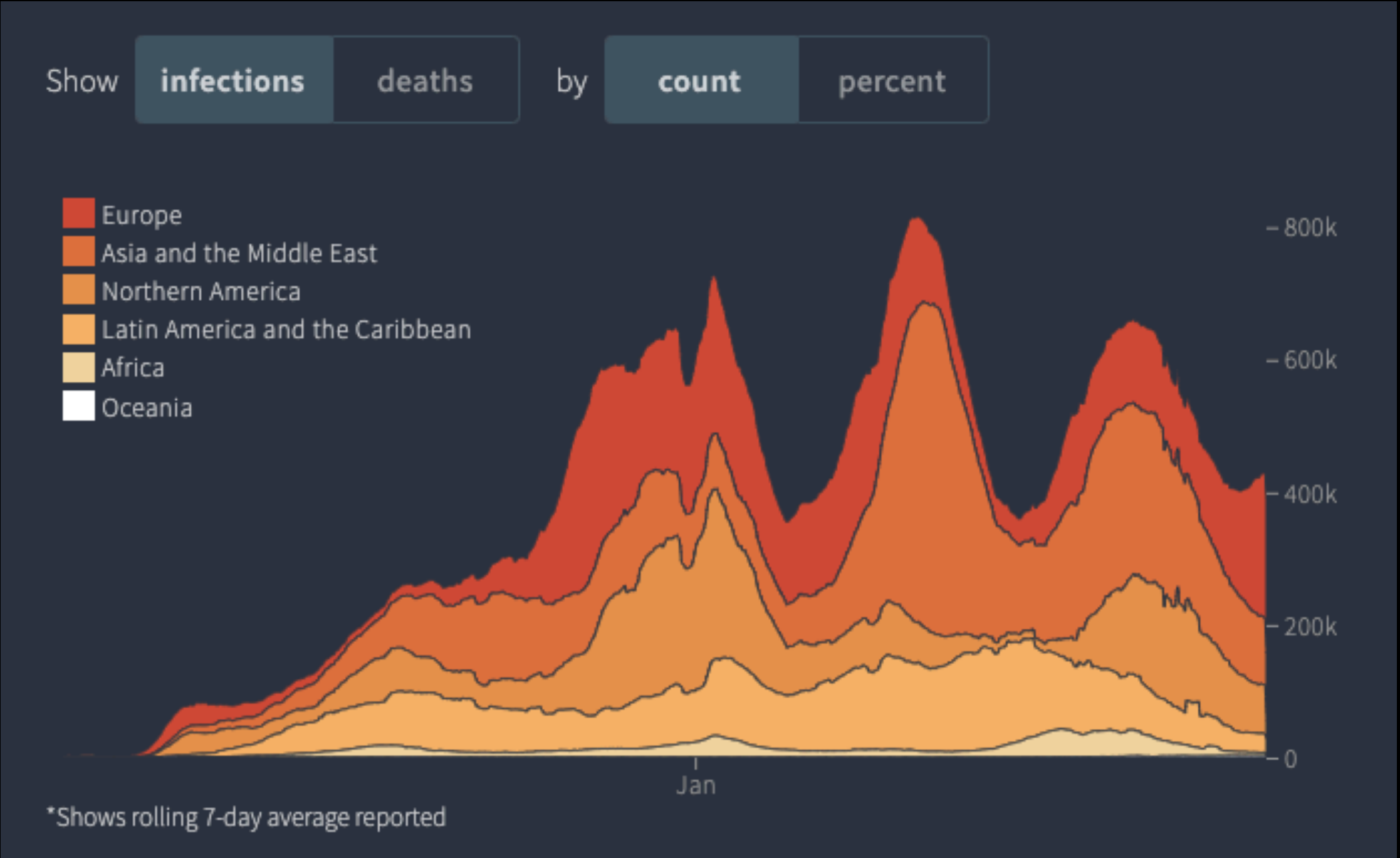
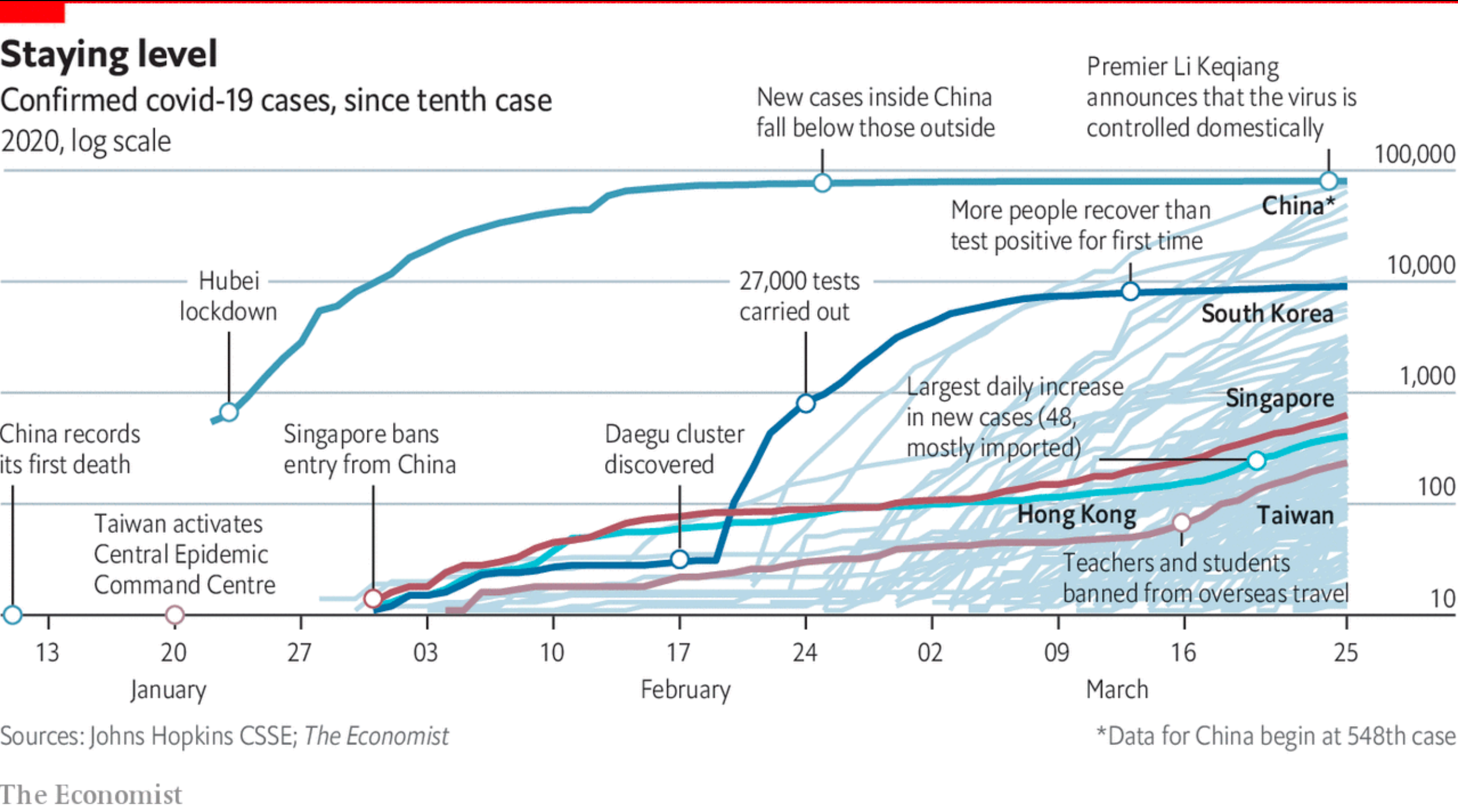
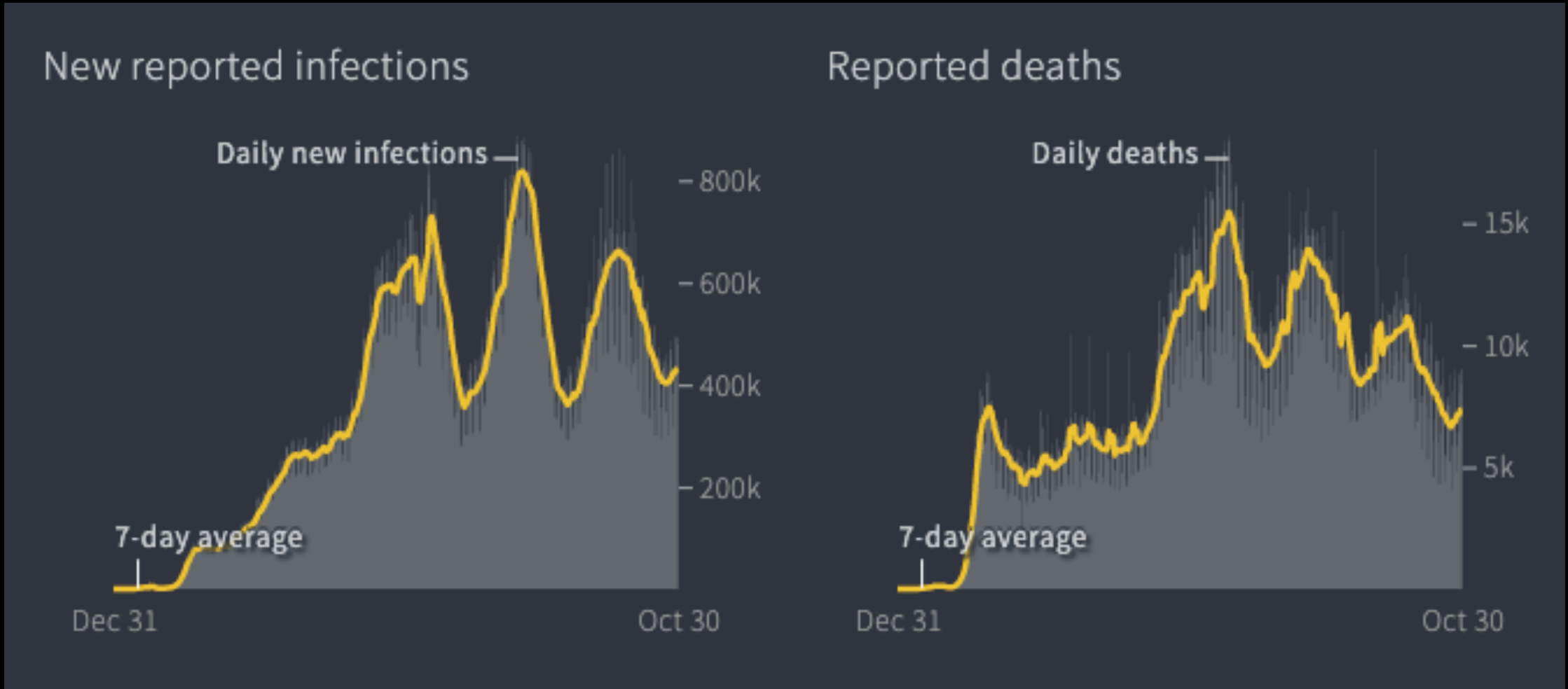
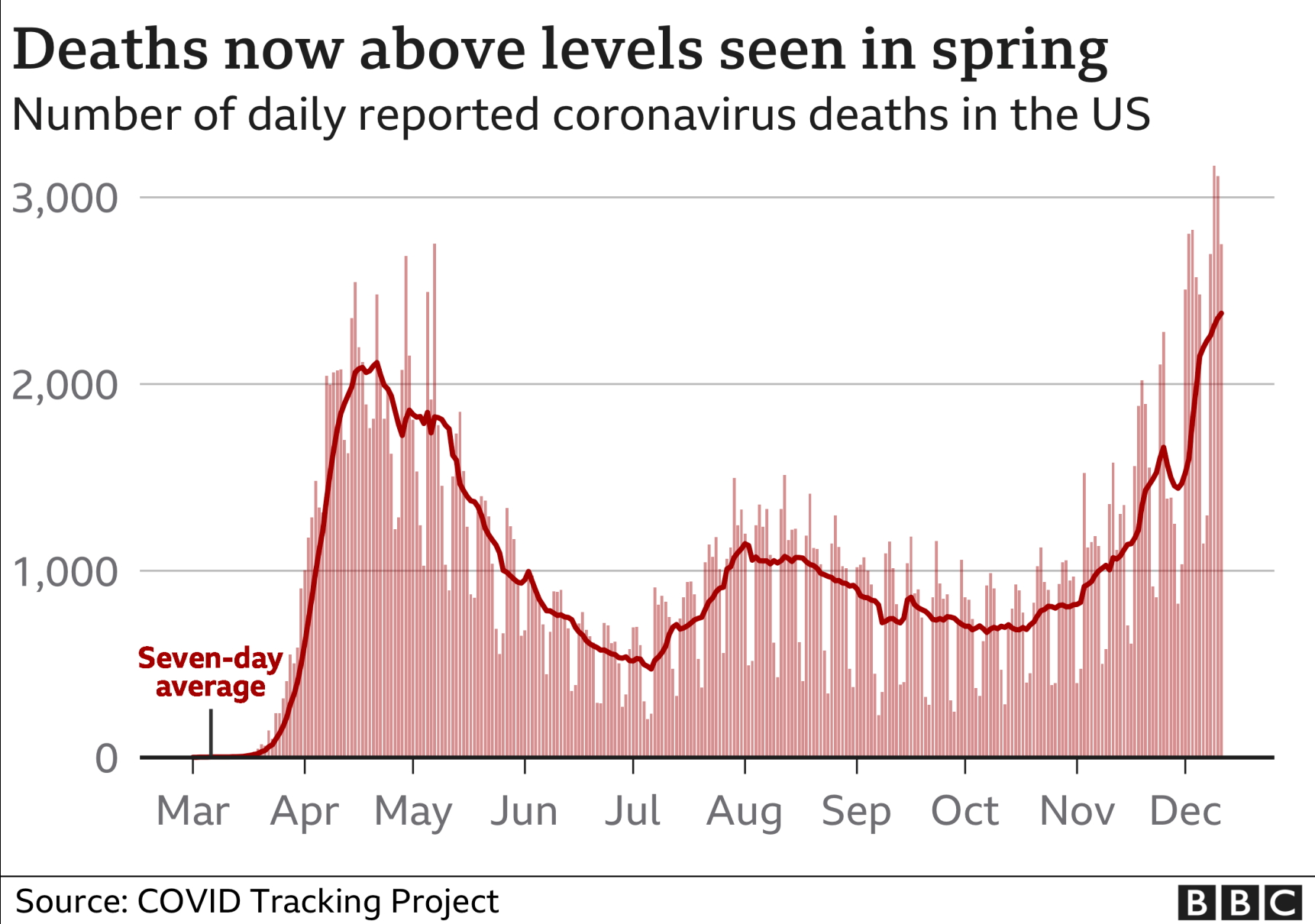
Evolving last week's LA Dome to this - a super detailed city model of [#Tokyo](#) including procedural street lamps, trees for parkland & even little aircraft warning lights on the tall buildings! Which do you prefer night or day?
[#3d](#) [#map](#) [#gis](#) [#design](#) [#animation](#) [#concept](#)



98 • 5 comments • 1,595 views

5. Increasing visual and data literacy

More content
means better
familiarity





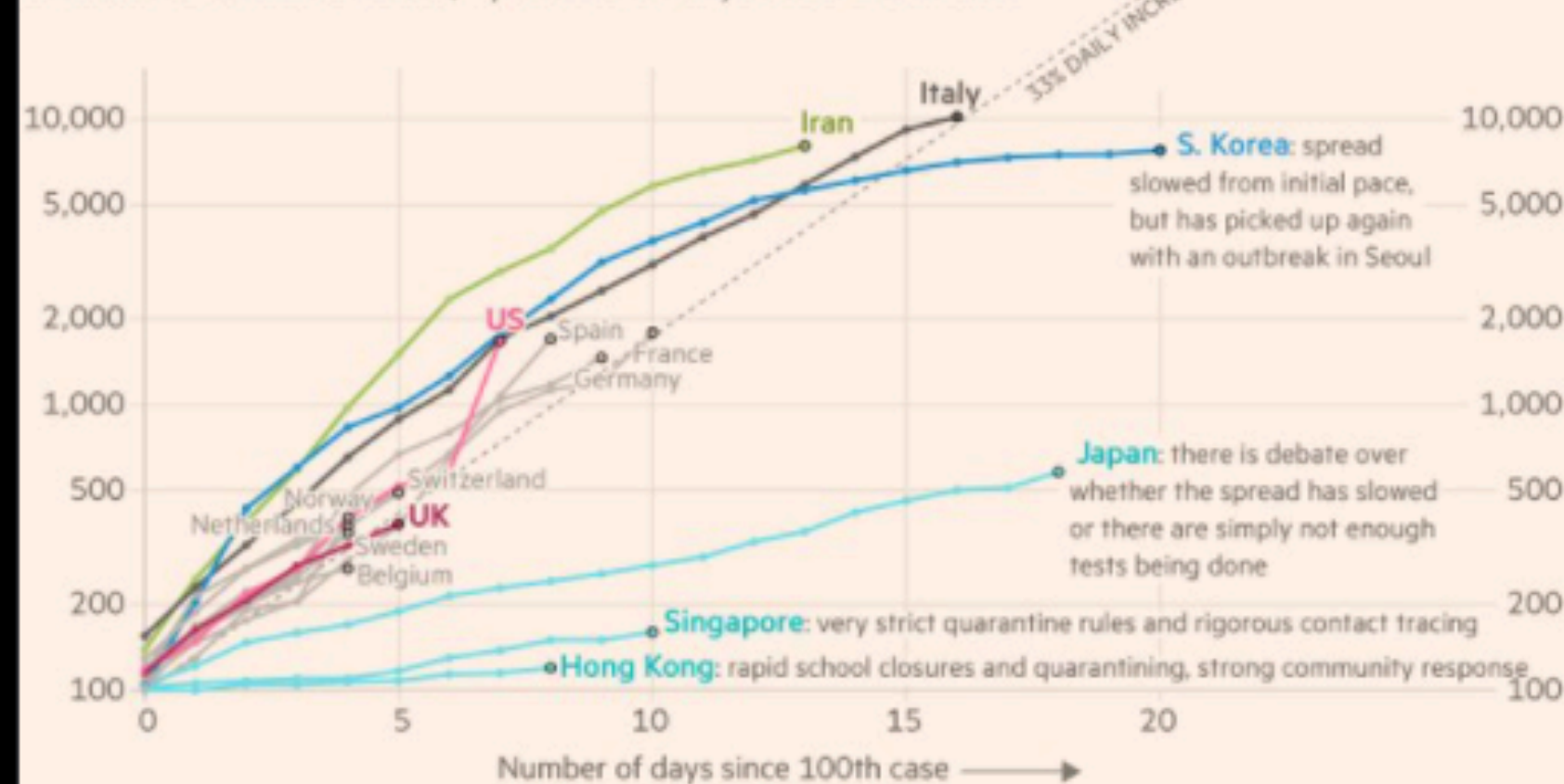
John Burn-Murdoch ✓
@jburnmurdoch

NEW on coronavirus: many western countries may soon face Italy's situation

Case numbers since outbreaks began in several countries have tracked a ~33% daily rise. This is as true for UK, France, Germany as Italy; the latter is simply further down the path [ft.com/content/a26fbf...](https://www.ft.com/content/a26fbf...)

Most western countries are on the same coronavirus trajectory. Hong Kong and Singapore have managed to slow the spread

Cumulative number of cases, by number of days since 100th case

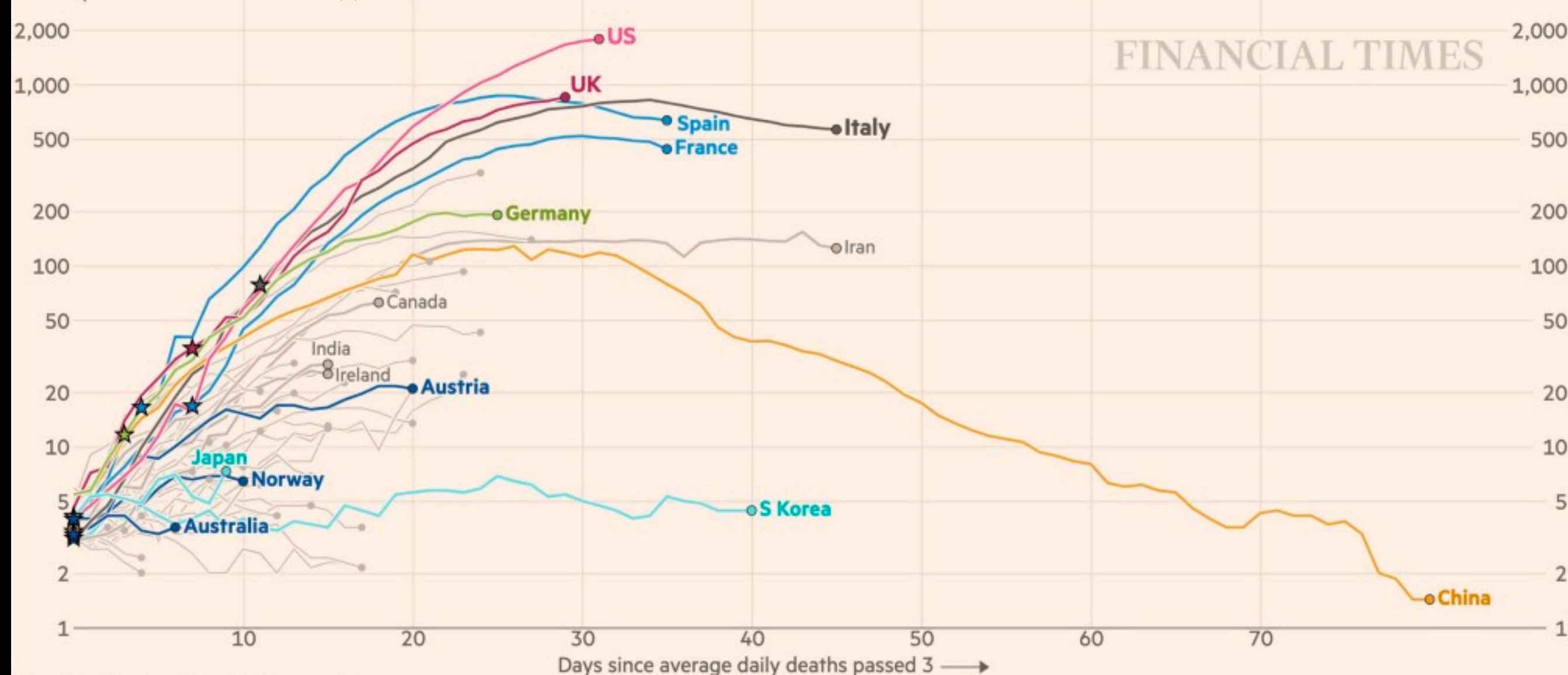


Source: FT analysis of Johns Hopkins University, CSSE
FT graphic: John Burn-Murdoch / @jburnmurdoch
© FT

7:19 pm · 11 Mar 2020 · Twitter Web App

Italy and Spain's daily death tolls are falling; in the UK and US daily deaths still trend upward

Daily deaths with coronavirus (7-day rolling average), by number of days since 3 daily deaths first recorded
Stars represent national lockdowns ★



FT graphic: John Burn-Murdoch / @jburnmurdoch
Source: FT analysis of European Centre for Disease Prevention and Control; FT research. Data updated April 13, 19:00 GMT
© FT

CORONAVIRUS TRAJECTORY TRACKER EXPLAINED

0:00 / 6:55

Video description Transcript

Explainer Data visualisation

Coronavirus trajectory tracker explained

The FT's senior data-visualisation journalist John Burn-Murdoch explains the must-see daily graphs, which show how cases and deaths are growing around the world

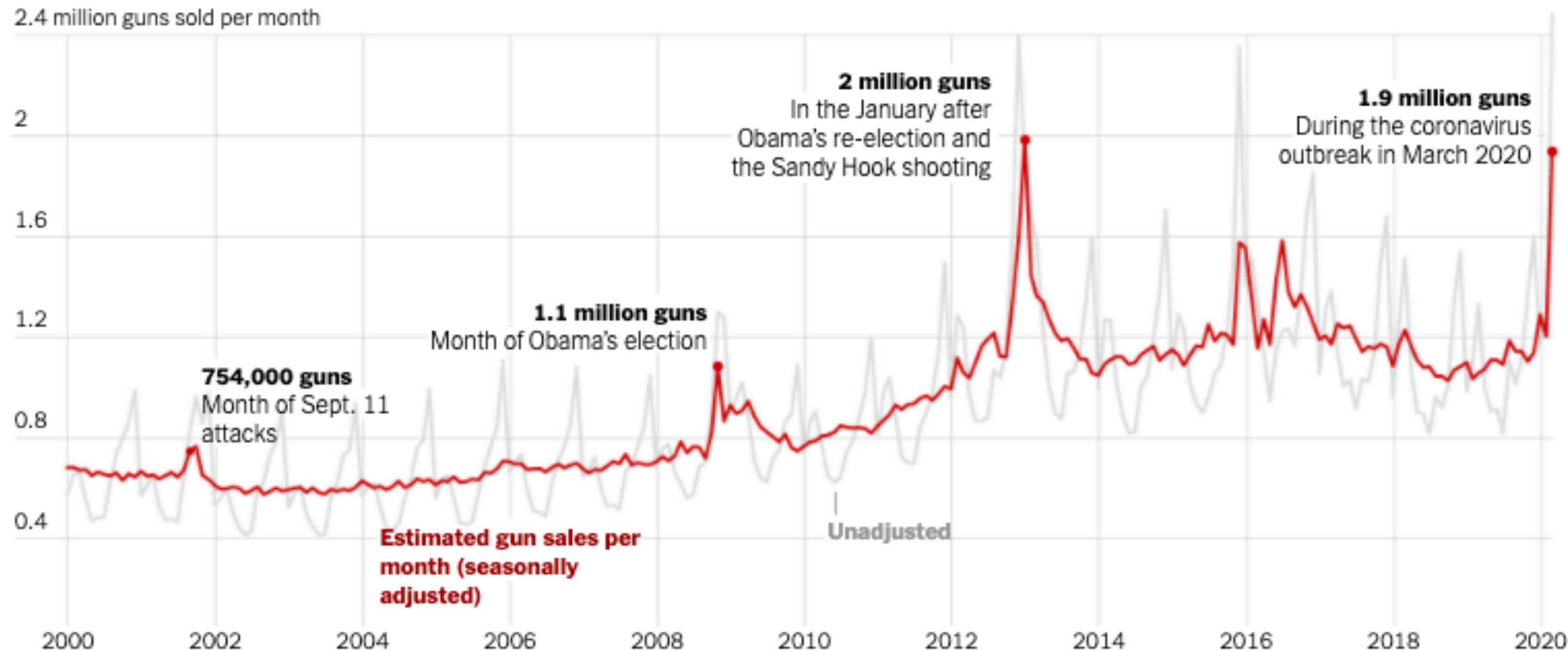
Twitter Facebook LinkedIn Save to myFT

March 30 2020 Produced by Tom Hannen

6. A slow and steady shift to stronger communication techniques

About 2 Million Guns Were Sold in the U.S. as Virus Fears Spread

By Keith Collins and David Yaffe-Bellany April 1, 2020



Source: The New York Times

“The annotation layer is
the most important thing
we do”

—Amanda Cox, Editor, Upshot

7. More data!

A DAY IN DATA

The exponential growth of data is undisputed, but the numbers behind this explosion - fuelled by internet of things and the use of connected devices - are hard to comprehend, particularly when looked at in the context of one day

500m

tweets are sent every day
Twitter



4PB

of data created by Facebook, including

350m photos

100m hours of video watch time

Facebook Research

DEMYSTIFYING DATA UNITS

From the more familiar 'bit' or 'megabyte', larger units of measurement are more frequently being used to explain the masses of data

Unit	Value	Size
b	bit	0 or 1
B	byte	8 bits
KB	kilobyte	1,000 bytes
MB	megabyte	1,000 ² bytes
GB	gigabyte	1,000 ³ bytes
TB	terabyte	1,000 ⁴ bytes
PB	petabyte	1,000 ⁵ bytes
EB	exabyte	1,000 ⁶ bytes
ZB	zettabyte	1,000 ⁷ bytes
YB	yottabyte	1,000 ⁸ bytes

*A lowercase "b" is used as an abbreviation for bits, while an uppercase "B" represents bytes.

65bn

messages sent over WhatsApp and two billion minutes of voice and video calls made

Facebook



294bn

billion emails are sent

Radicati Group

320bn

emails to be sent each day by 2021

306bn

emails to be sent each day by 2020

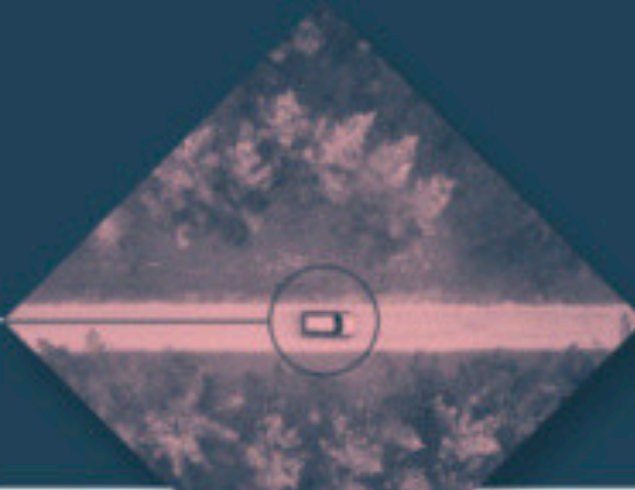
3.9bn

people use emails

4TB

of data produced by a connected car

Intel



ACCUMULATED DIGITAL UNIVERSE OF DATA

4.4ZB

2013

44ZB

2020

PwC

Searches made a day

5bn

Searches made a day from Google

3.5bn

Smart Insights



463EB

of data will be created every day by 2025

ISC

95m

photos and videos are shared on Instagram

Instagram Business



28PB

to be generated from wearable devices by 2020

Statista



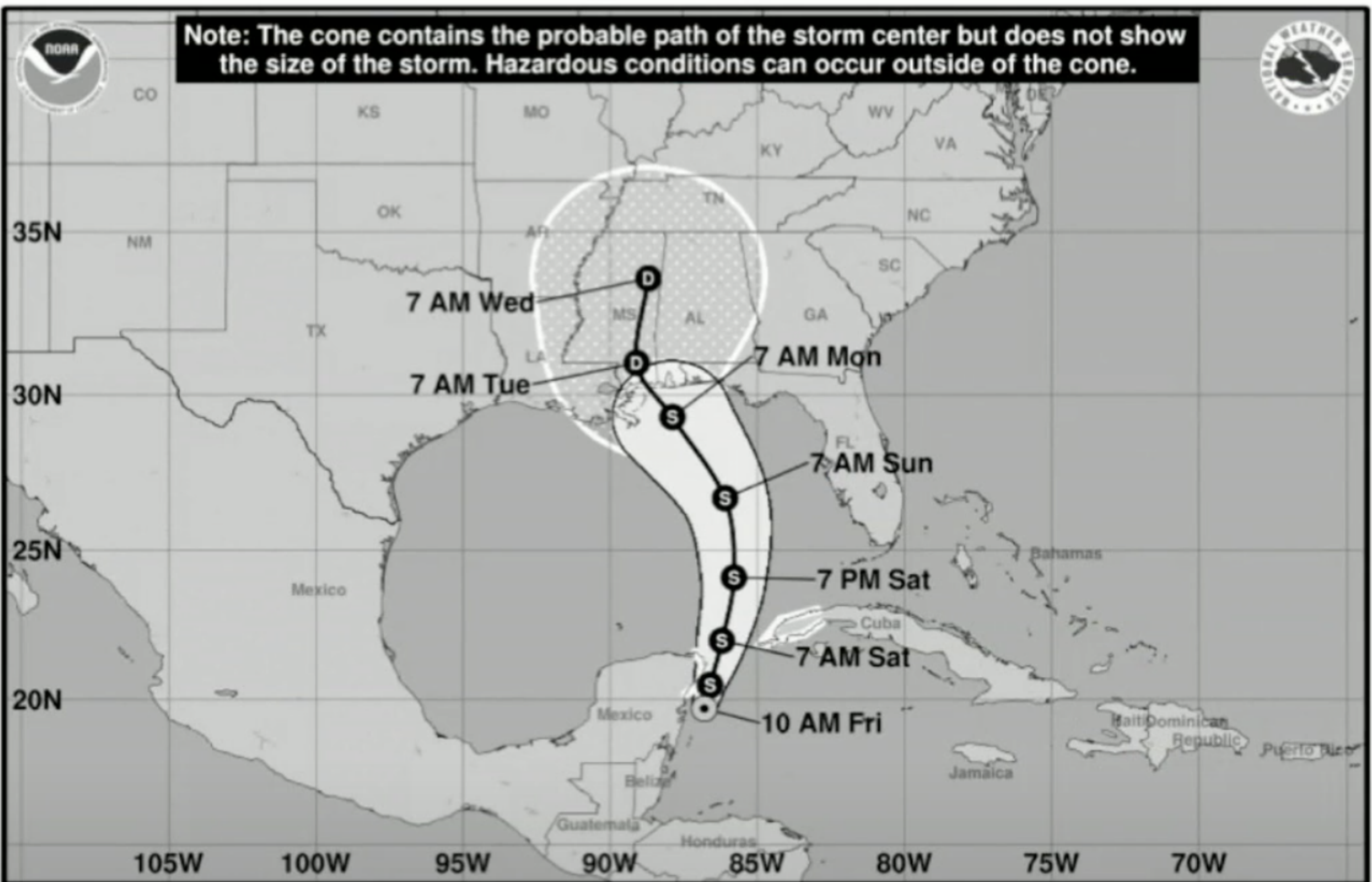
RACONTEUR

But we have
much further to
go...

1. We still are pretty
bad at communicating
uncertainty



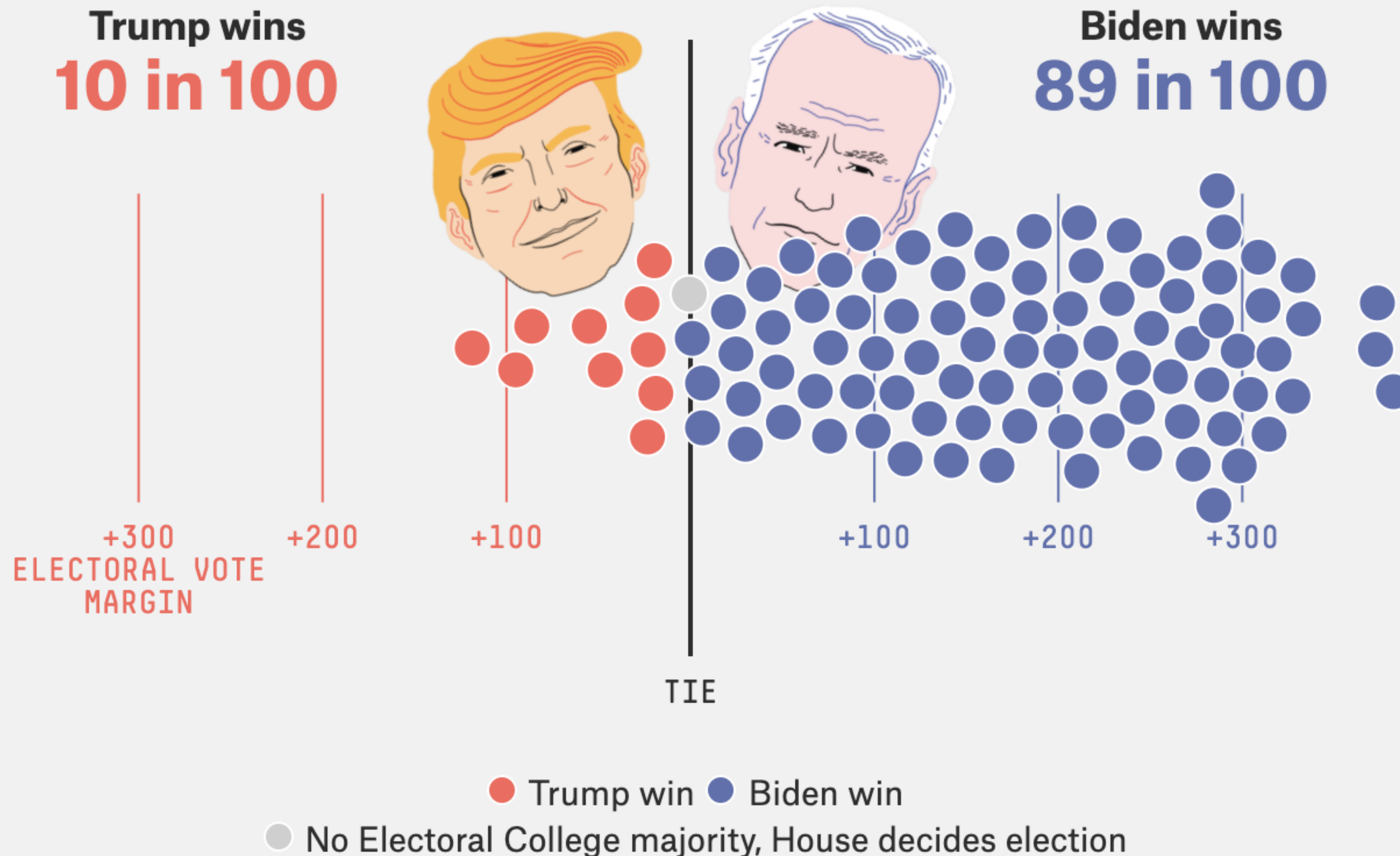
Note: The cone contains the probable path of the storm center but does not show the size of the storm. Hazardous conditions can occur outside of the cone.



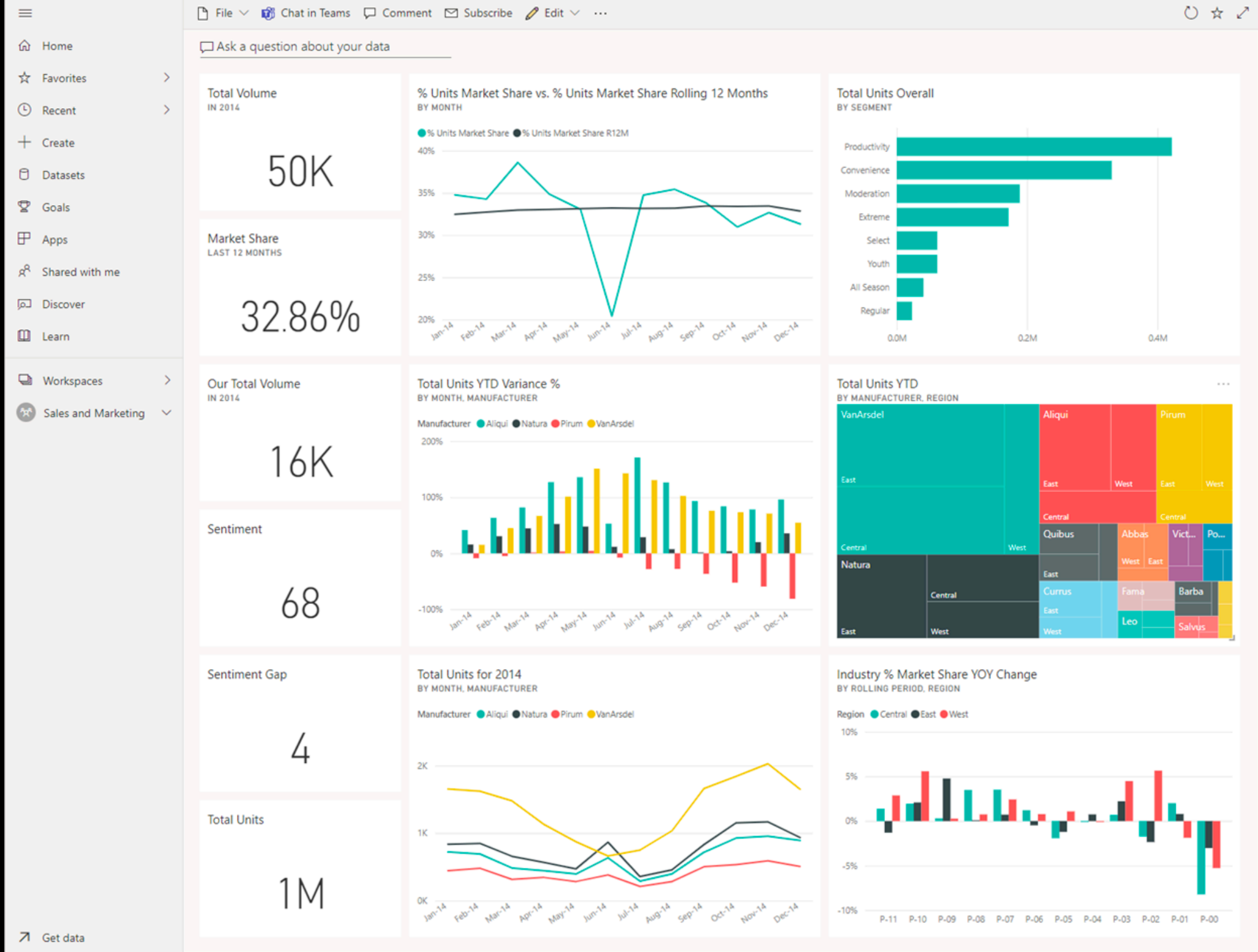
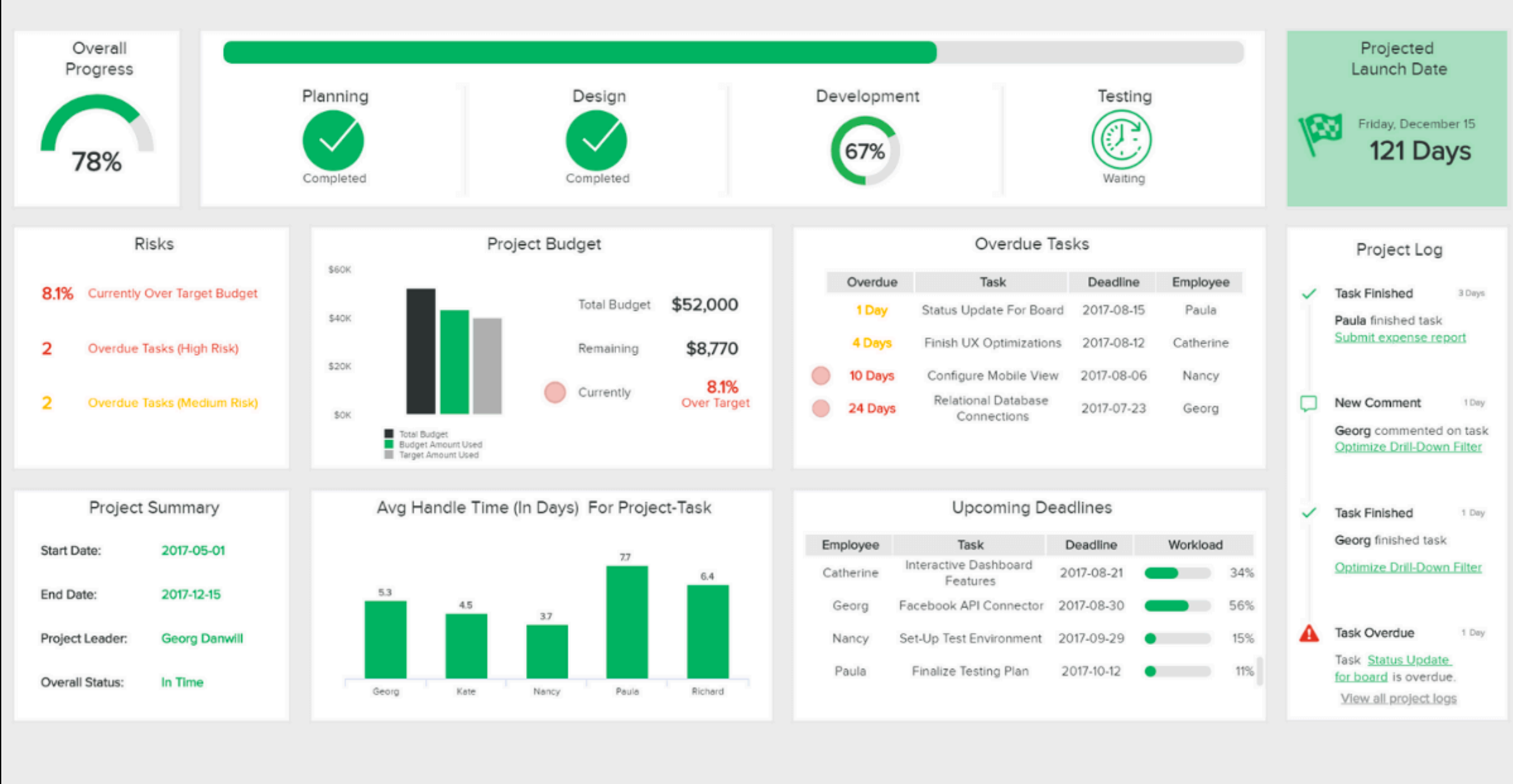
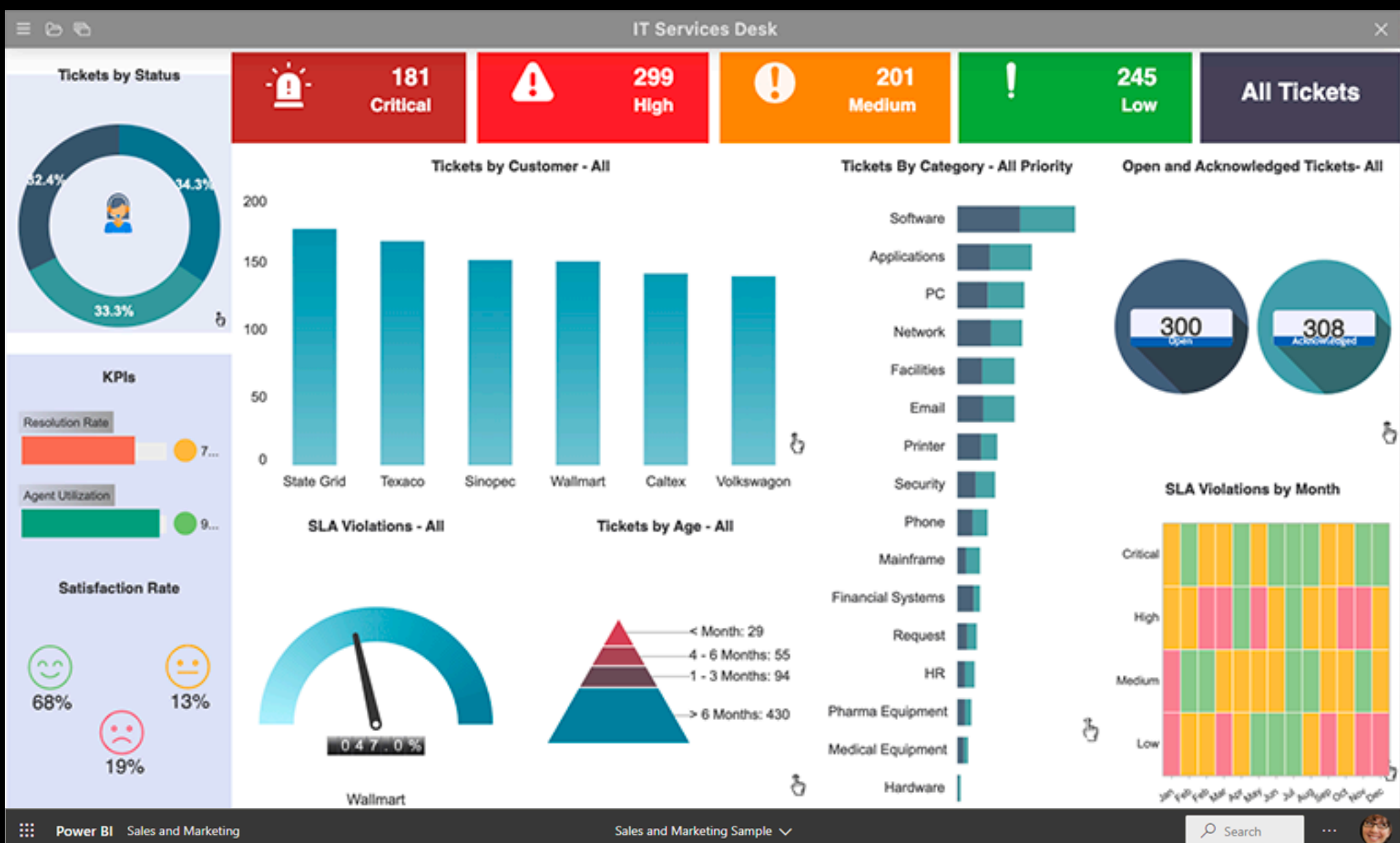
Subtropical Storm Alberto Friday May 25, 2018 10 AM CDT Advisory 1 NWS National Hurricane Center	Current information: ● Center location 19.7 N 86.8 W Maximum sustained wind 40 mph Movement NNE at 6 mph	Forecast positions: ● Tropical Cyclone ○ Post/Potential TC Sustained winds: D < 39 mph S 39-73 mph H 74-110 mph M > 110 mph
Potential track area: △ Day 1-3 ▨ Day 4-5	Watches: ■ Hurricane ■ Trop Storm	Warnings: ■ Hurricane ■ Trop Storm

Biden is *favored* to win the election

We simulate the election 40,000 times to see who wins most often. The sample of 100 outcomes below gives you a good idea of the range of scenarios our model thinks is possible.

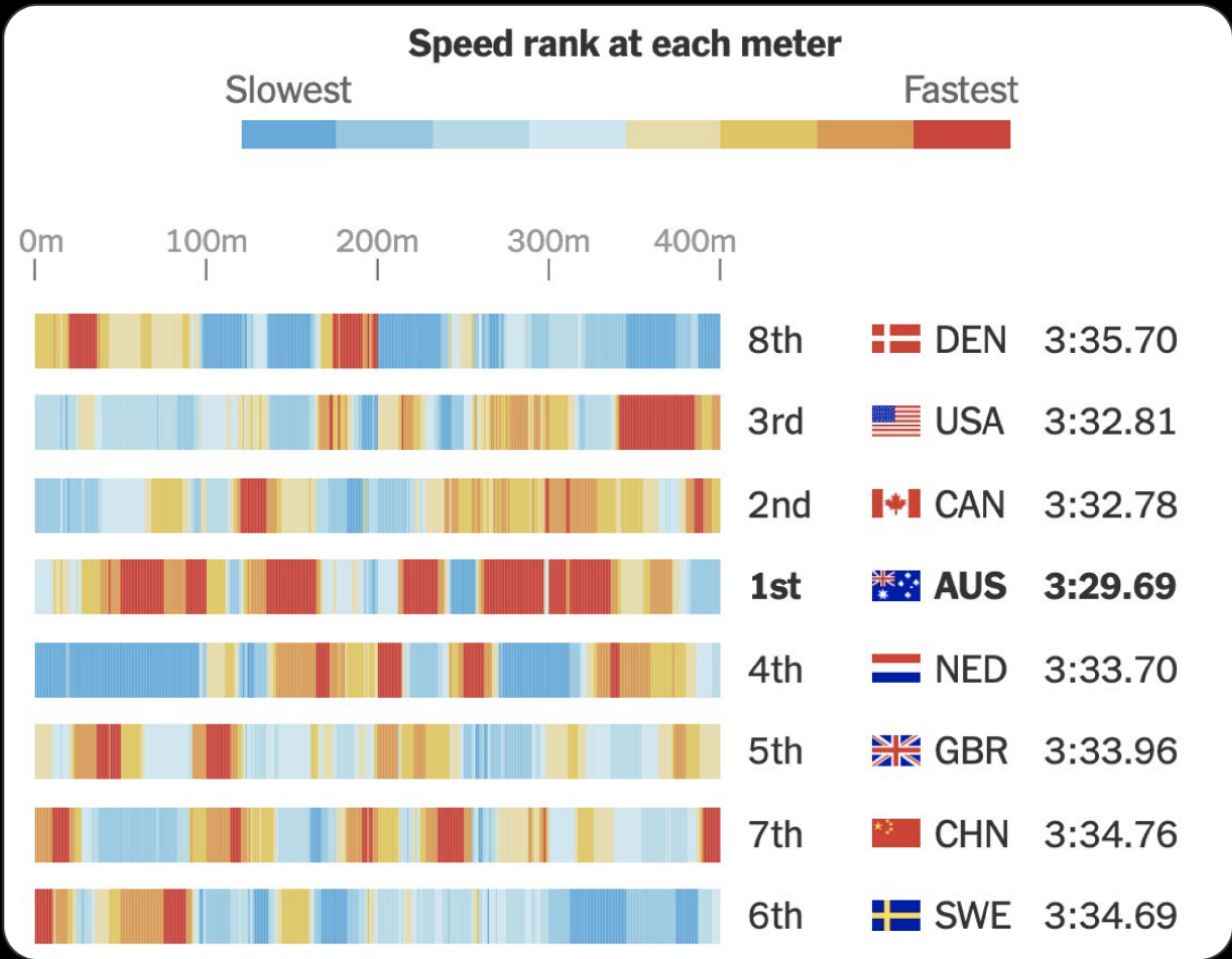


2. More data also means
more noise to filter
through

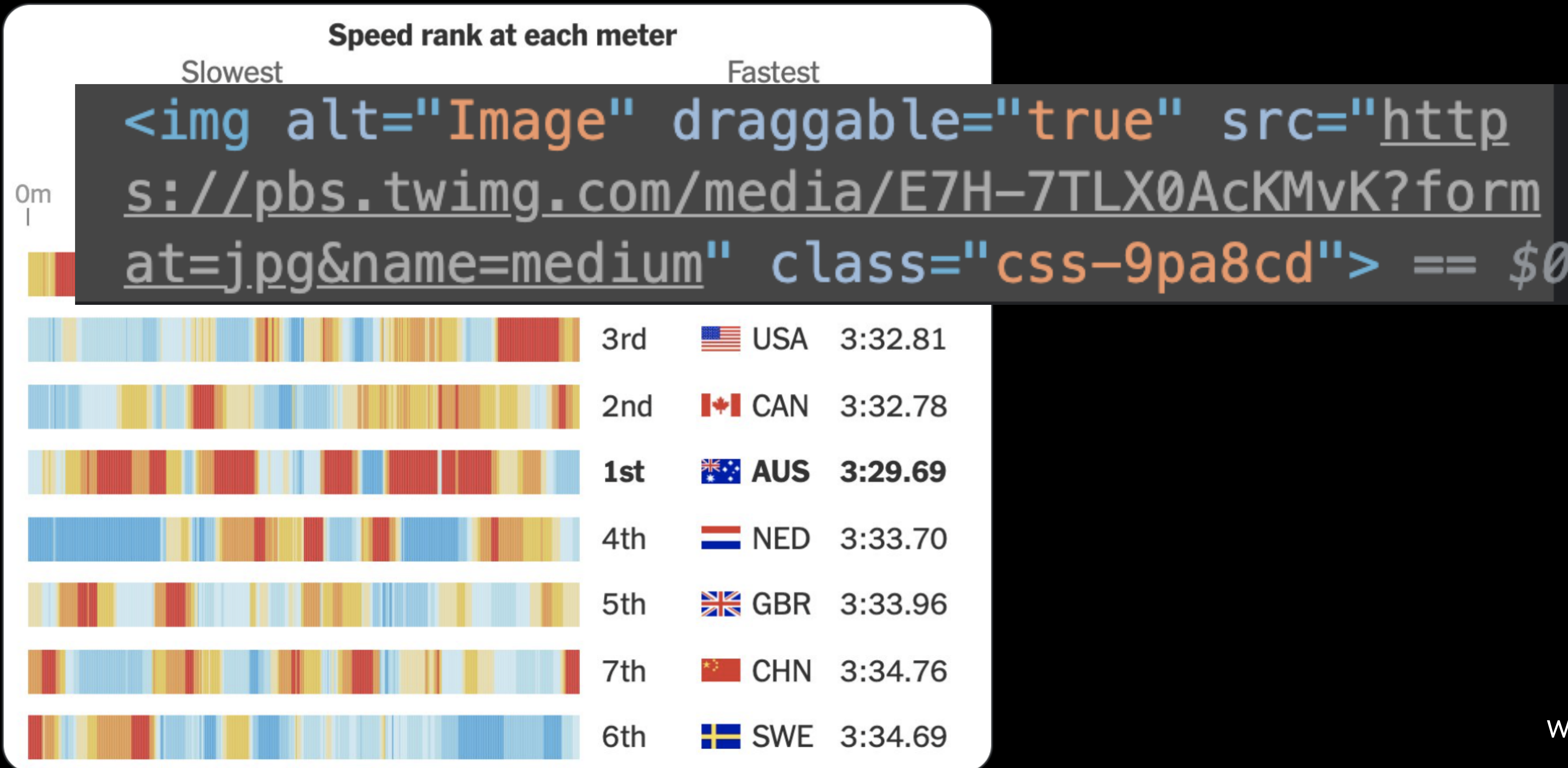


3. A lot of dataviz is
still not “accessible”

See how Australia dominated the women's 4x100 freestyle relay, swimming some of the fastest speeds across all four legs. [nytimes.com/interactive/20...](https://www.nytimes.com/interactive/2020/07/23/sports/olympics/4x100-freestyle-relay.html)



See how Australia dominated the women's 4x100 freestyle relay, swimming some of the fastest speeds across all four legs. [nytimes.com/interactive/20...](https://www.nytimes.com/interactive/2016/08/12/sports/olympics/4x100-freestyle-relay-women.html)

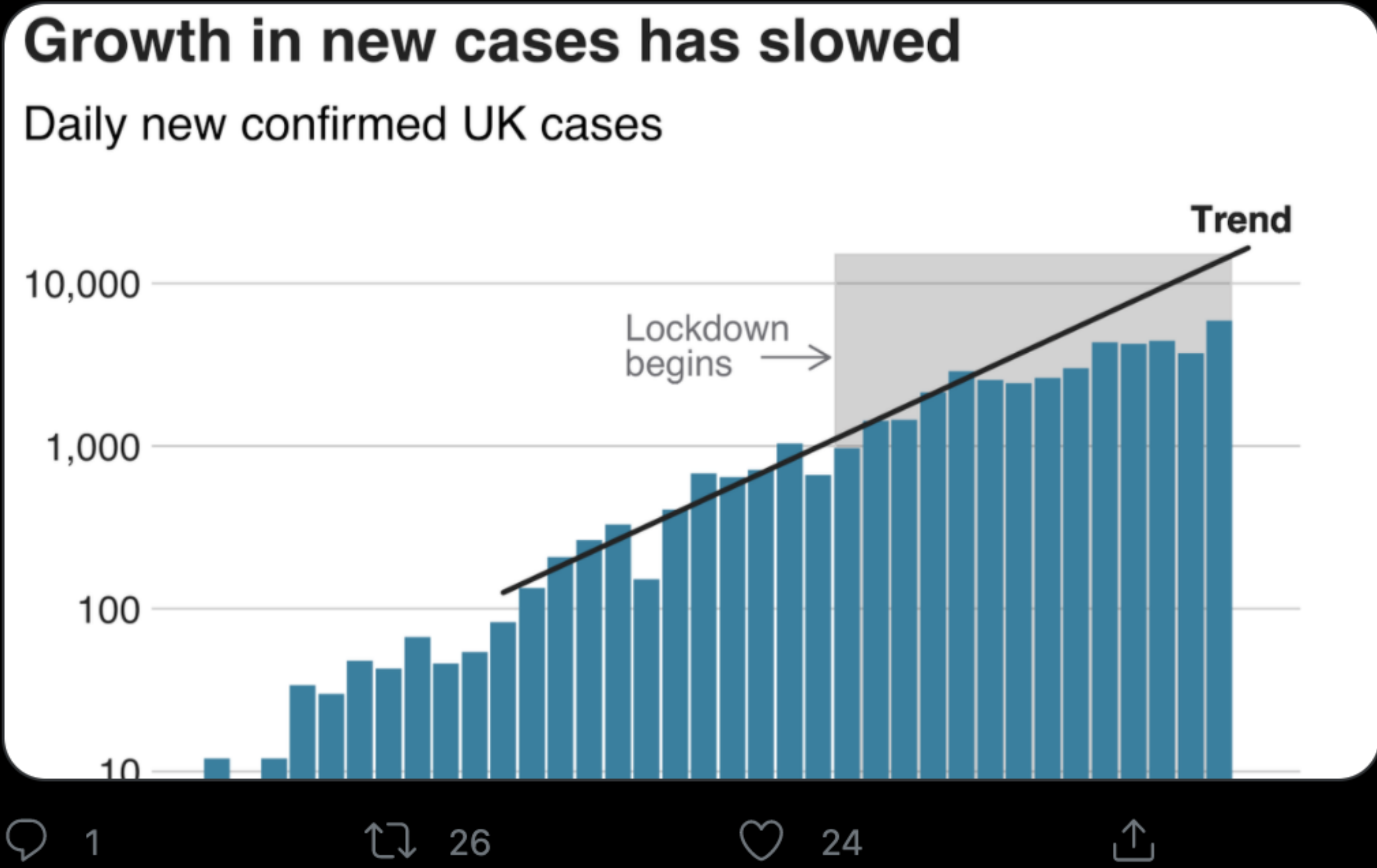




BBC News Graphics @BBCNewsGraphics · Apr 6, 2020

The UK's [#coronavirus](#) lockdown measures are not yet showing up in the death data but other figures do give clues to their effectiveness.

bbc.co.uk/news/health-52...

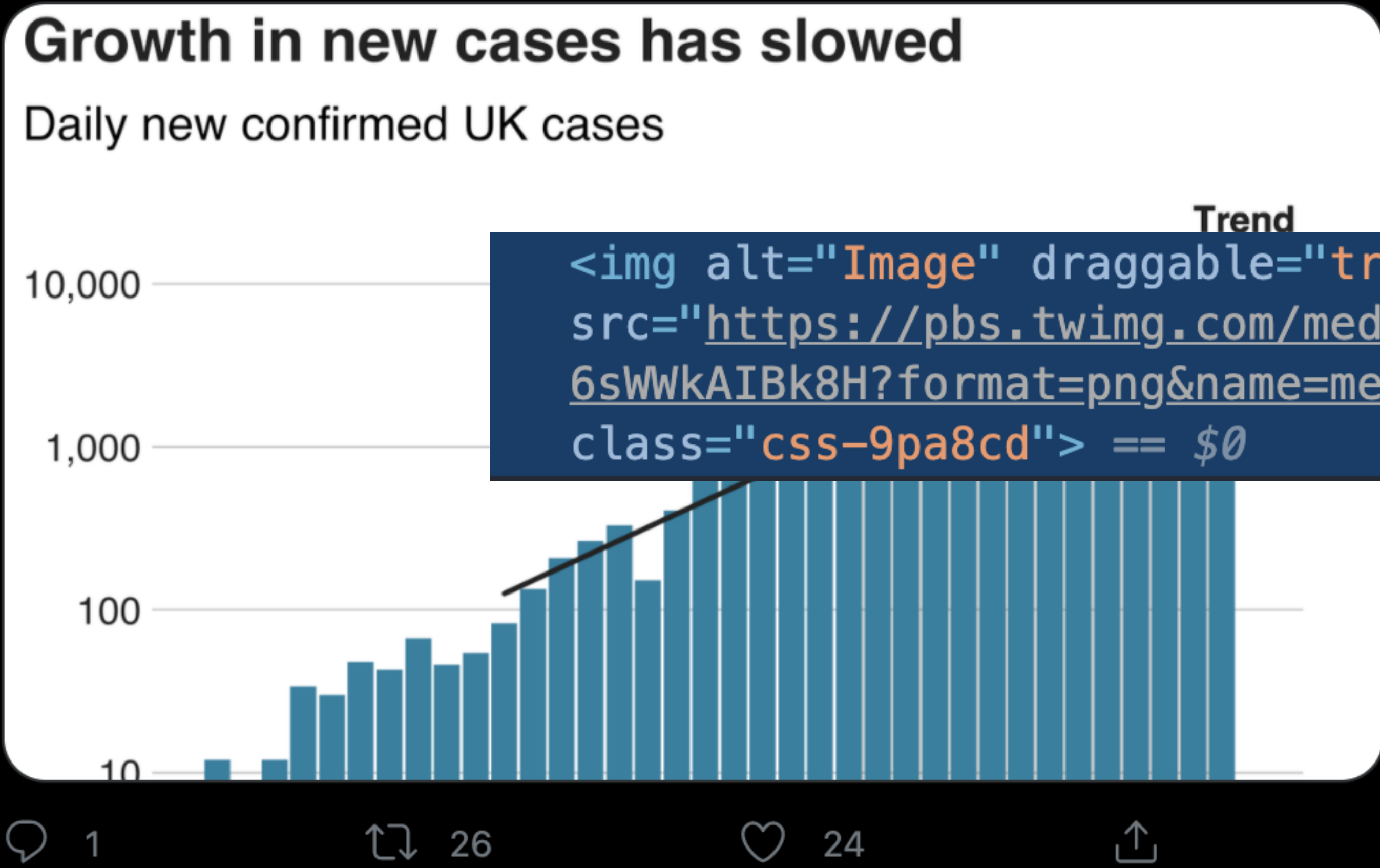




BBC News Graphics @BBCNewsGraphics · Apr 6, 2020

The UK's [#coronavirus](#) lockdown measures are not yet showing up in the death data but other figures do give clues to their effectiveness.

[bbc.co.uk/news/health-52...](#)



Monday, 23. November 2020

How (not) to make accessible data visualizations, illustrated by the US presidential election.

Dataviz

Accessibility

Design

Development

4. Our tools still don't help
us achieve “data variety”

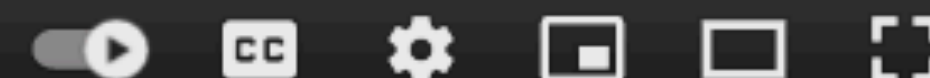
CHARTING A PATH FORWARD

from **Designers** to **Decision Makers**

from **Specification** to **Exploration**

from **Design Variation** to **Data Variation**

9:43 / 10:10



"The Future of Data Visualization" - Jeffrey Heer (Strata + Hadoop 2015)

152,421 views • Feb 21, 2015

1.2K 22 SHARE SAVE ...

**“Show data variation,
not design variation”**

– Edward Tufte

The humans still have to
innovate with the
questions...

How can our tools help us ask better questions?

**Finding the data
and the story**

**Communicating
the story**

Finding the data and the story

No amount of design can
save a bad story.

Thank you

Gurman Bhatia
@gurmanbhatia
www.gurmanbhatia.com