

Time

Session 13

PMAP 8921: Data Visualization with R
Andrew Young School of Policy Studies
Summer 2024

Plan for today

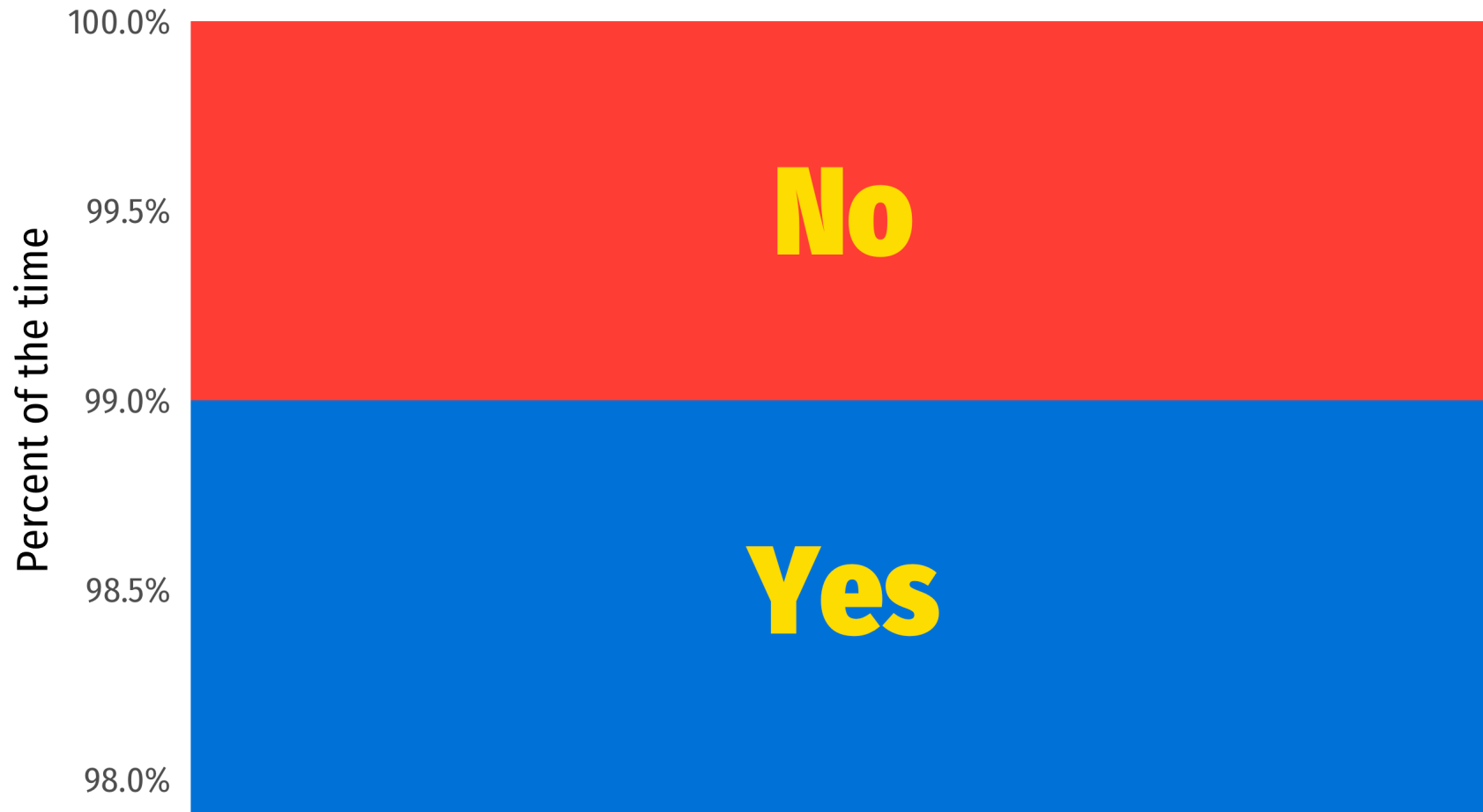
Axis issues

Visualizing time

**Starting, ending,
and decomposing time**

Axis issues

Is truncating the y-axis misleading?



Don't be too extreme!

It is actually more legal to truncate the y-axis than you might think!

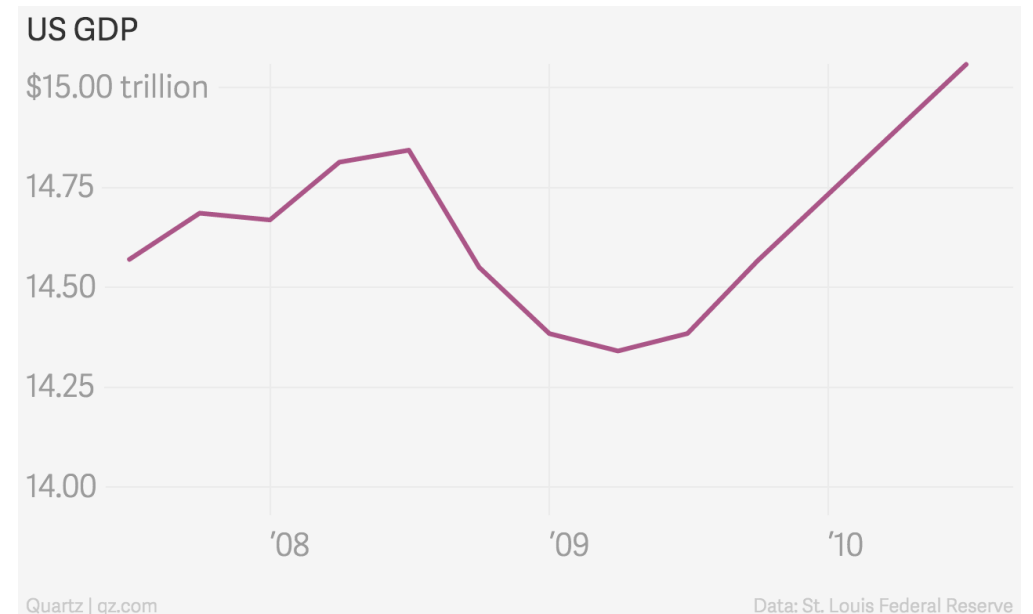
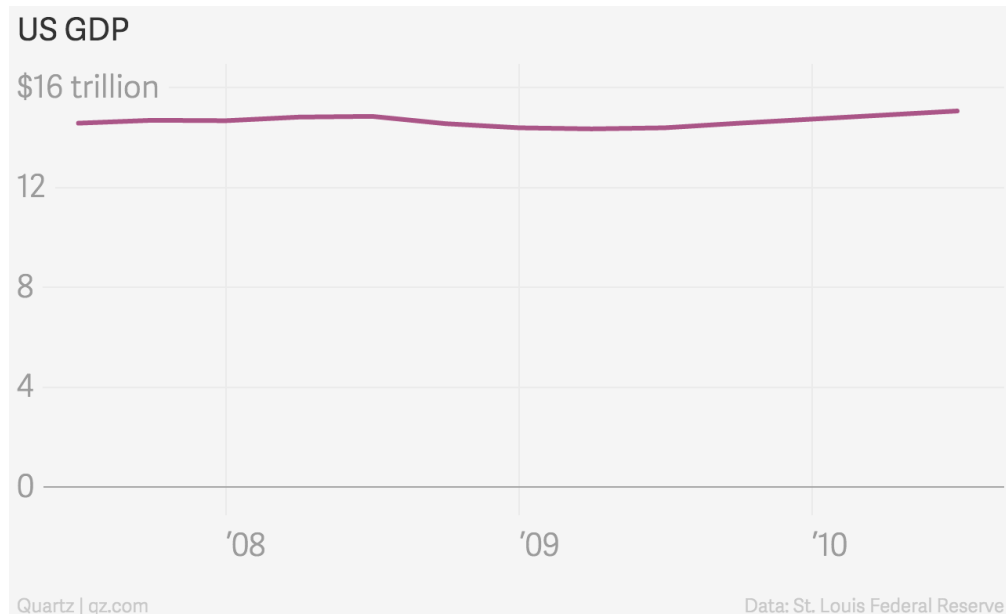
When small movements matter

When the scale itself is distorted

When zero values are impossible

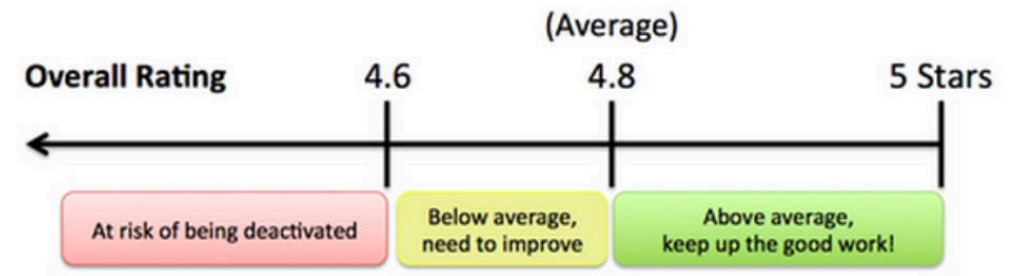
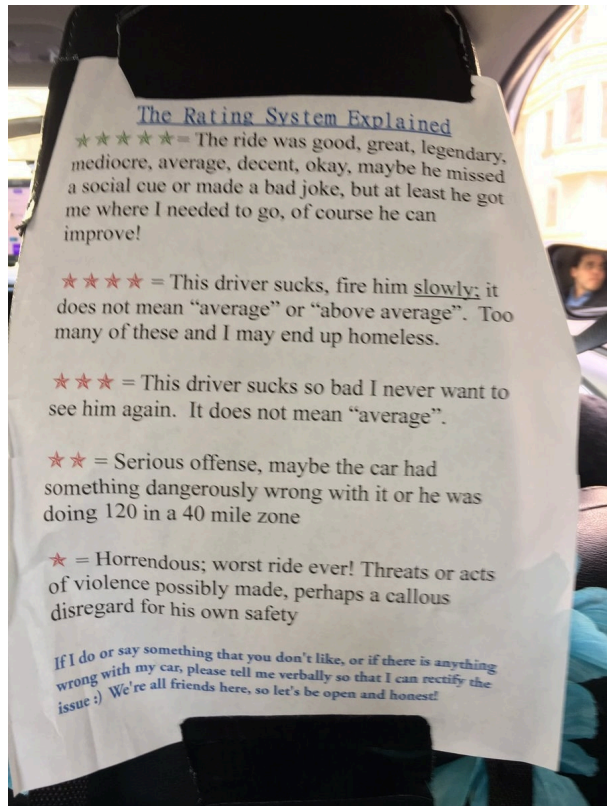
When is it okay to truncate?

When small movements matter



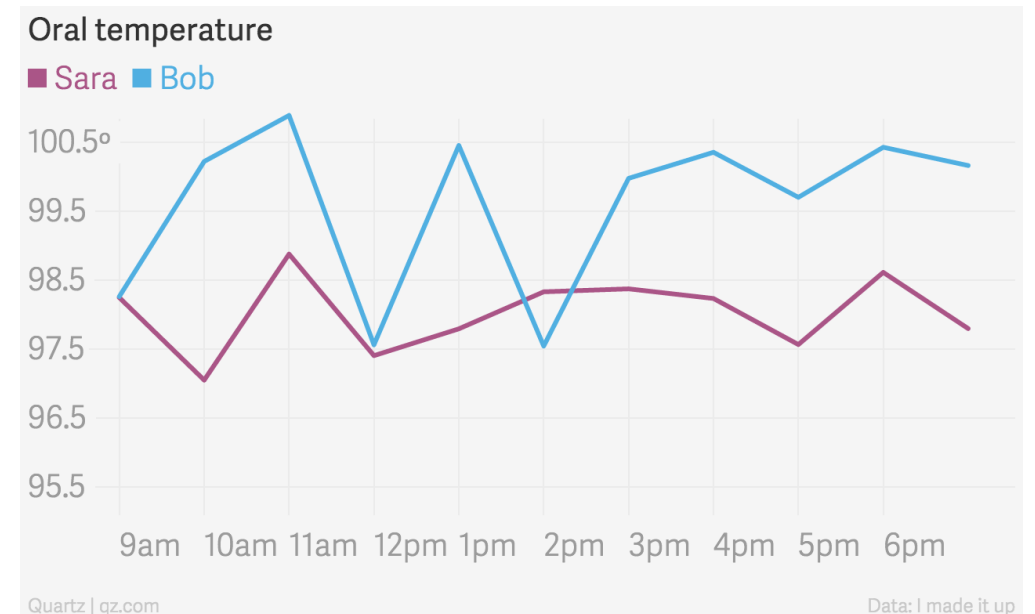
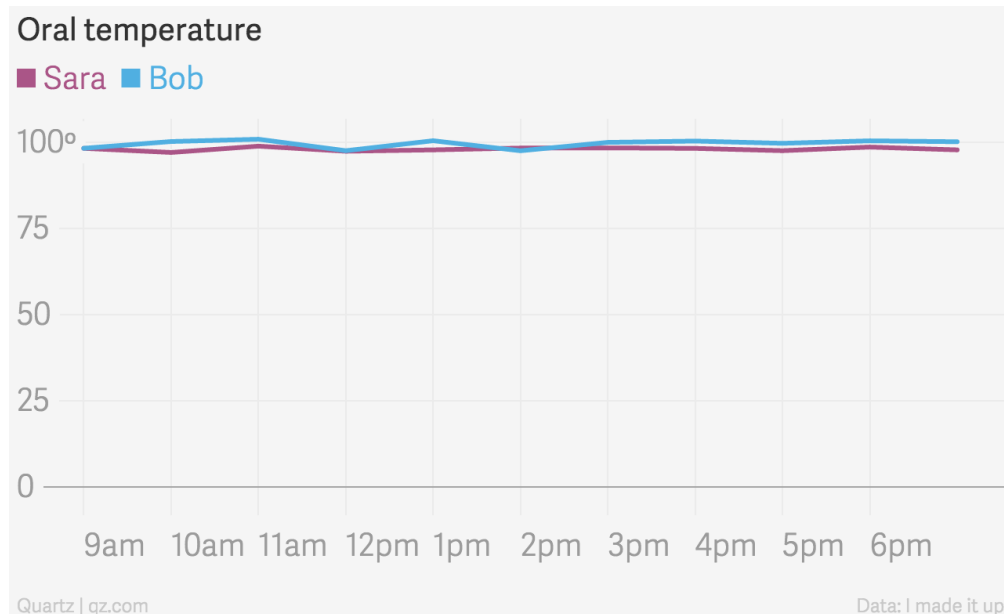
When is it okay to truncate?

When the scale itself is distorted

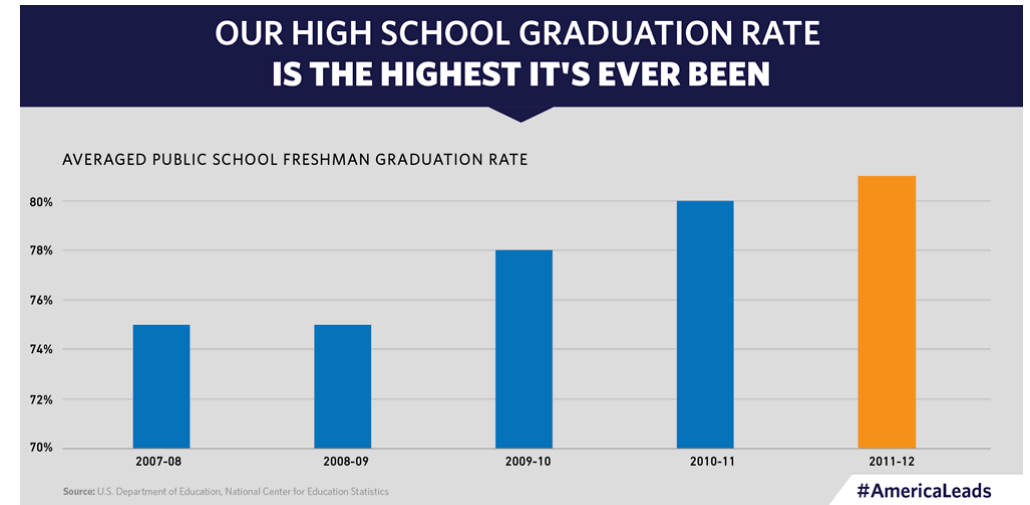


When is it okay to truncate?

When zero values are impossible



Never on bar charts

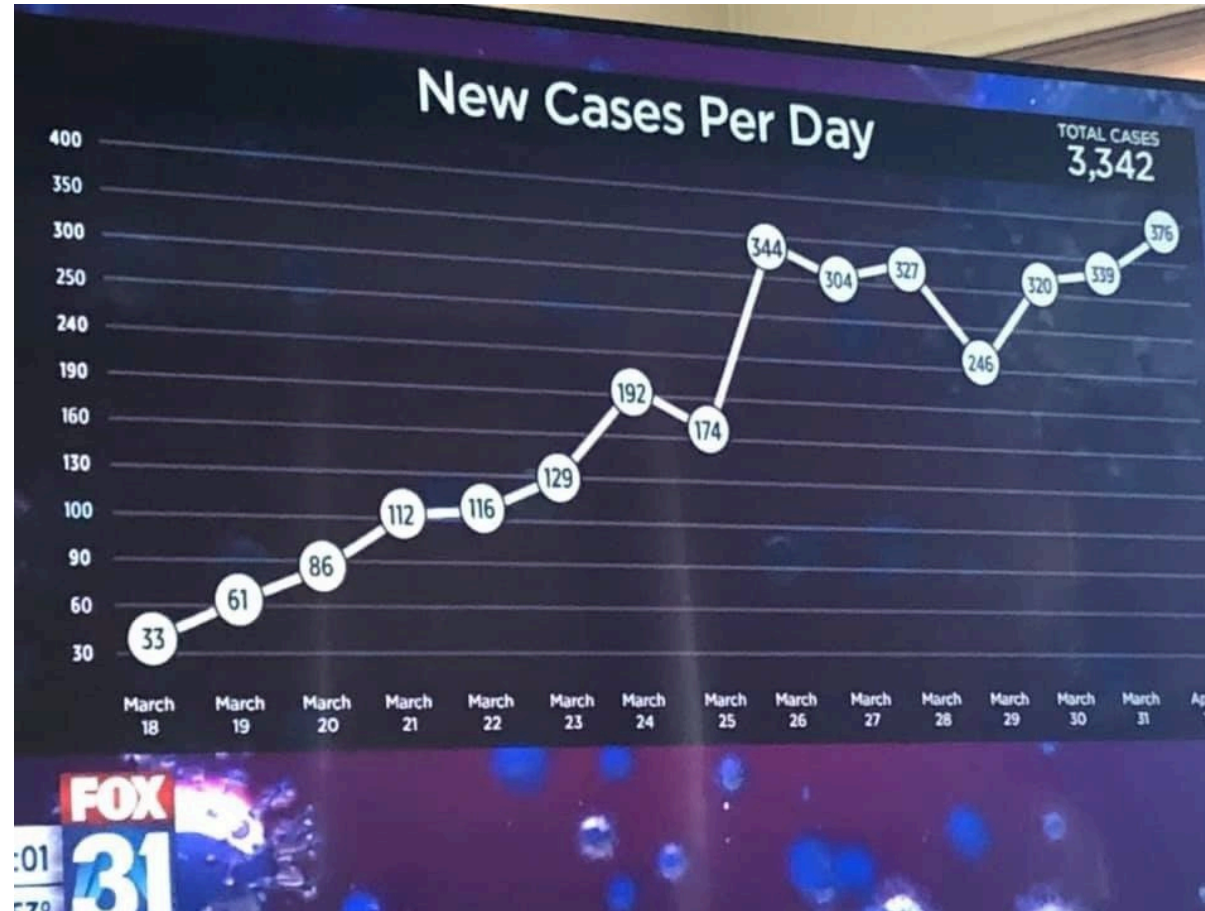


Zero is okay too!

Just because you don't *have to* start at 0 doesn't mean you should *never* start at 0

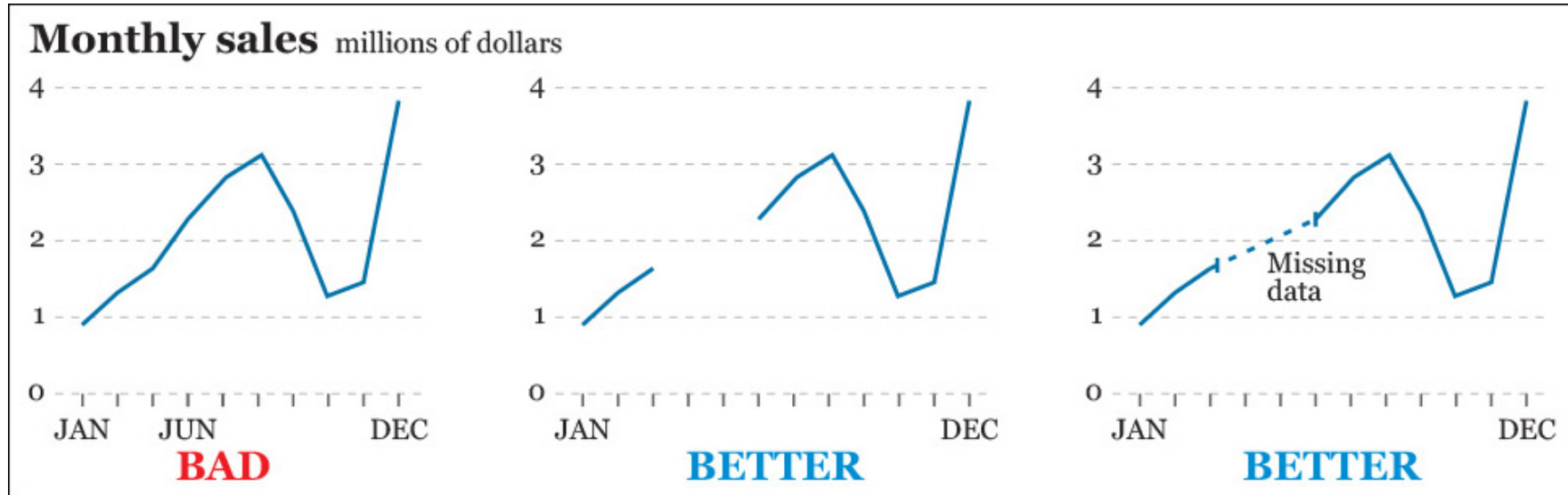
It's often a good idea!

Keep axis scales consistent

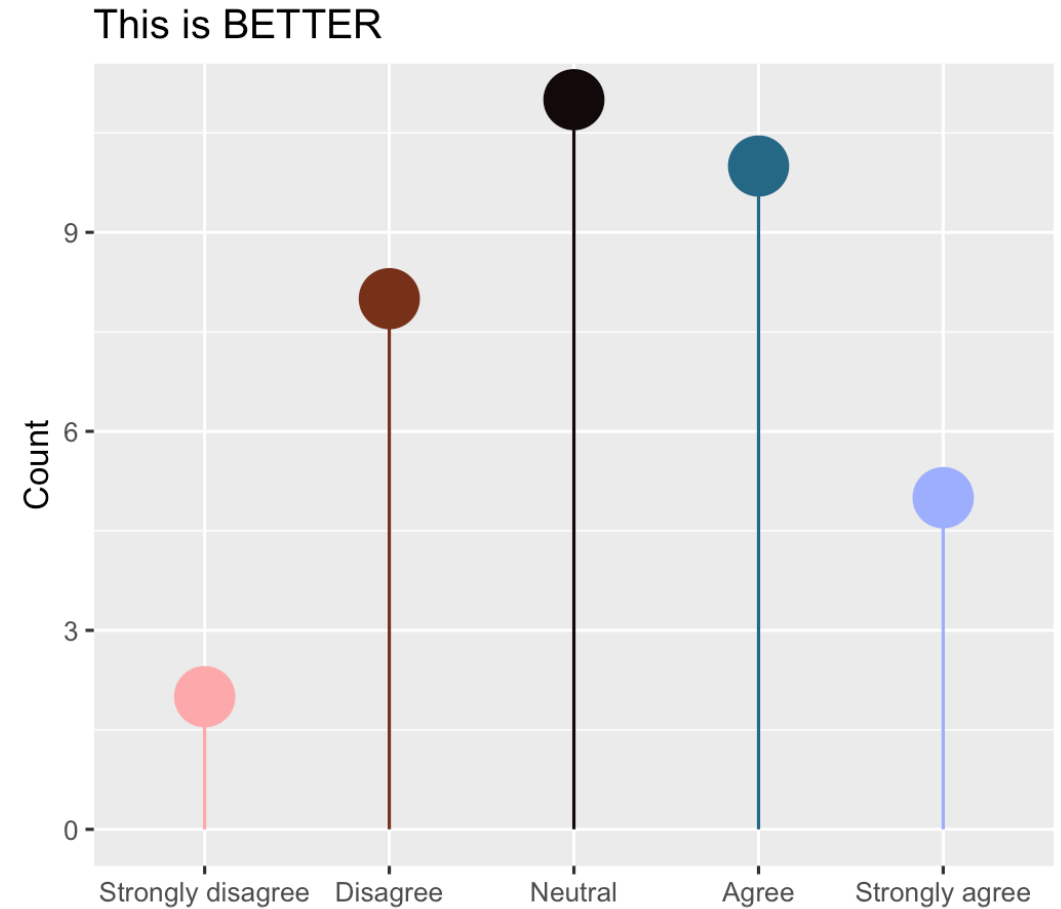
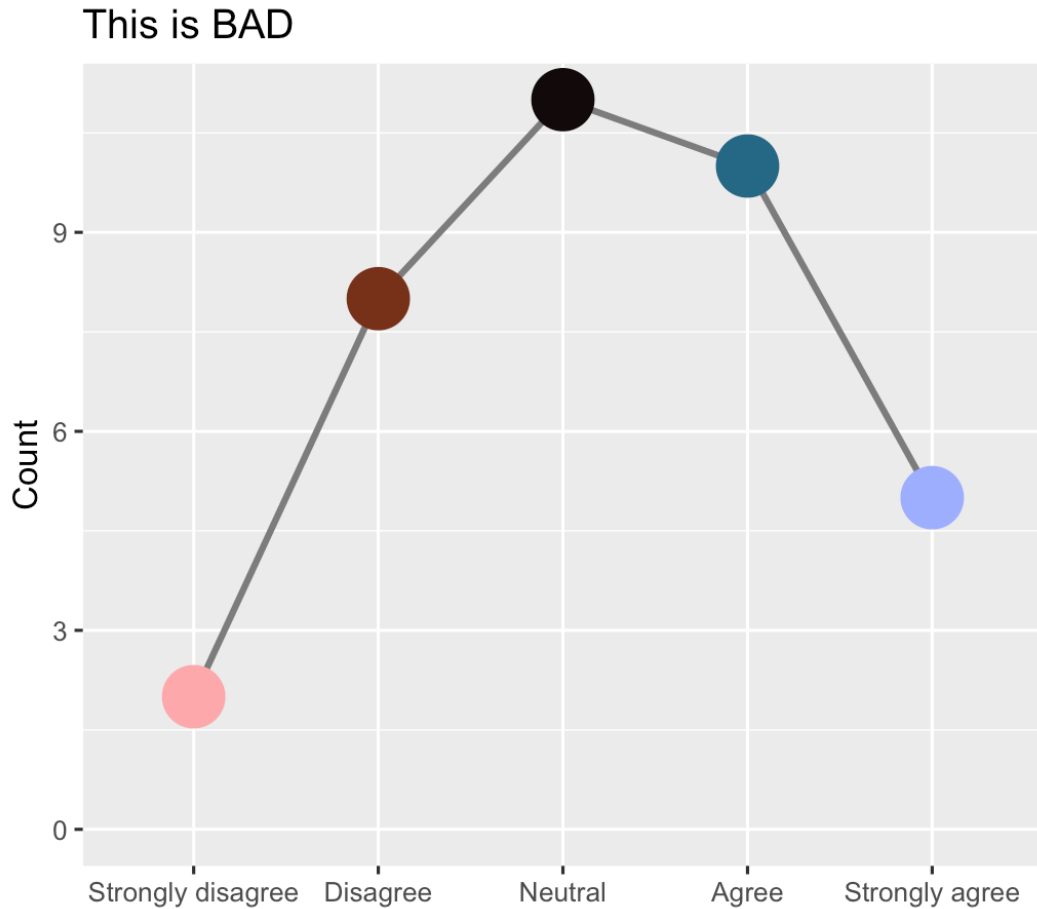


FOX affiliate in Colorado reporting on COVID-19 cases

Keep axis scales consistent



Don't impute across categories



Visualizing time

Showing changes over time

Time is just a variable that can be mapped to an aesthetic

Can be used as `x`, `y`, `color`, `fill`, `facet`, and even animation

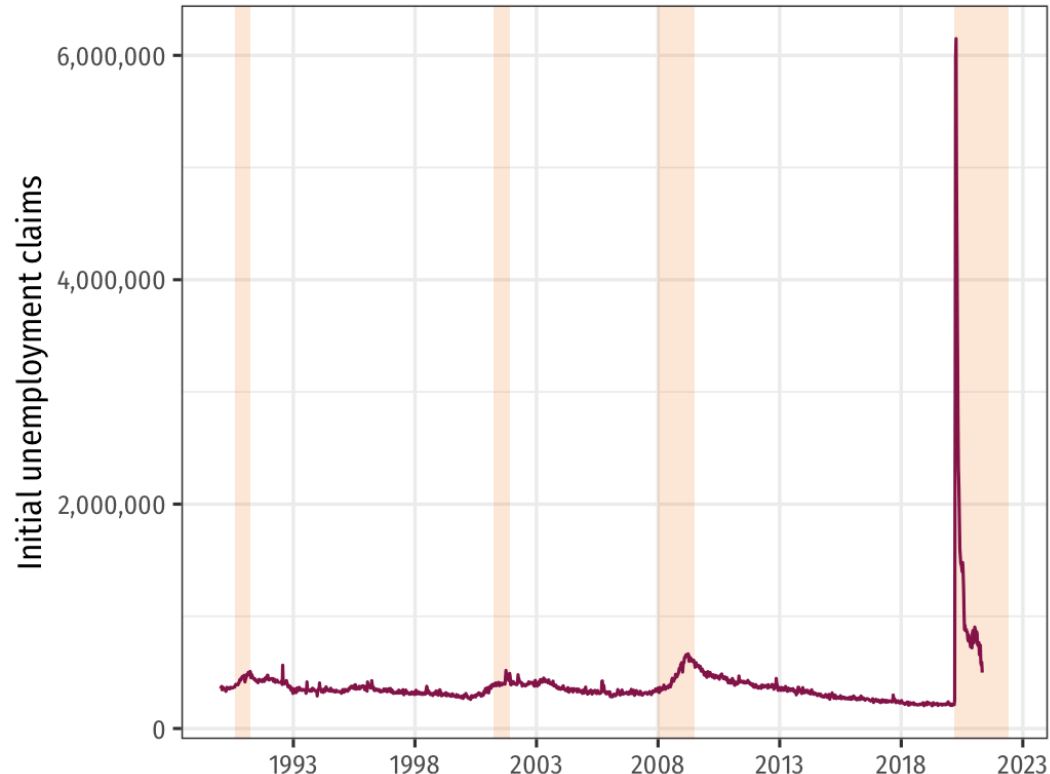
Can use all sorts of `geom` S:
lines, columns, points, heatmaps, densities, maps, etc.

In general, follow reading conventions to show time progression:



Time on x-axis + geom_line/col()

HOLY CRAP



Source: Initial weekly unemployment claims (ICSA); FRED Recessions highlighted in orange



Front page of the New York Times, May 9, 2020

Time on x-axis + geom_tile()

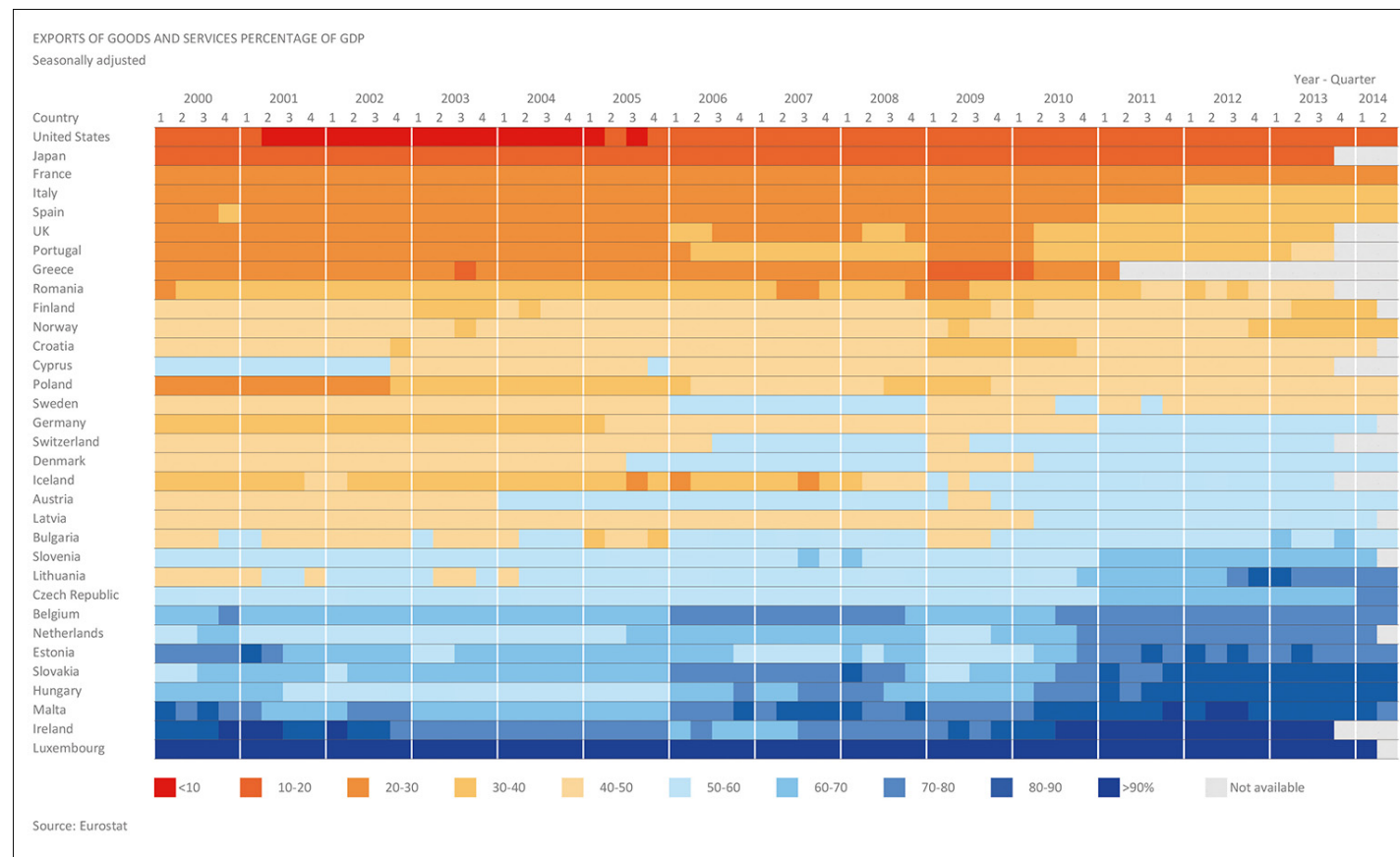
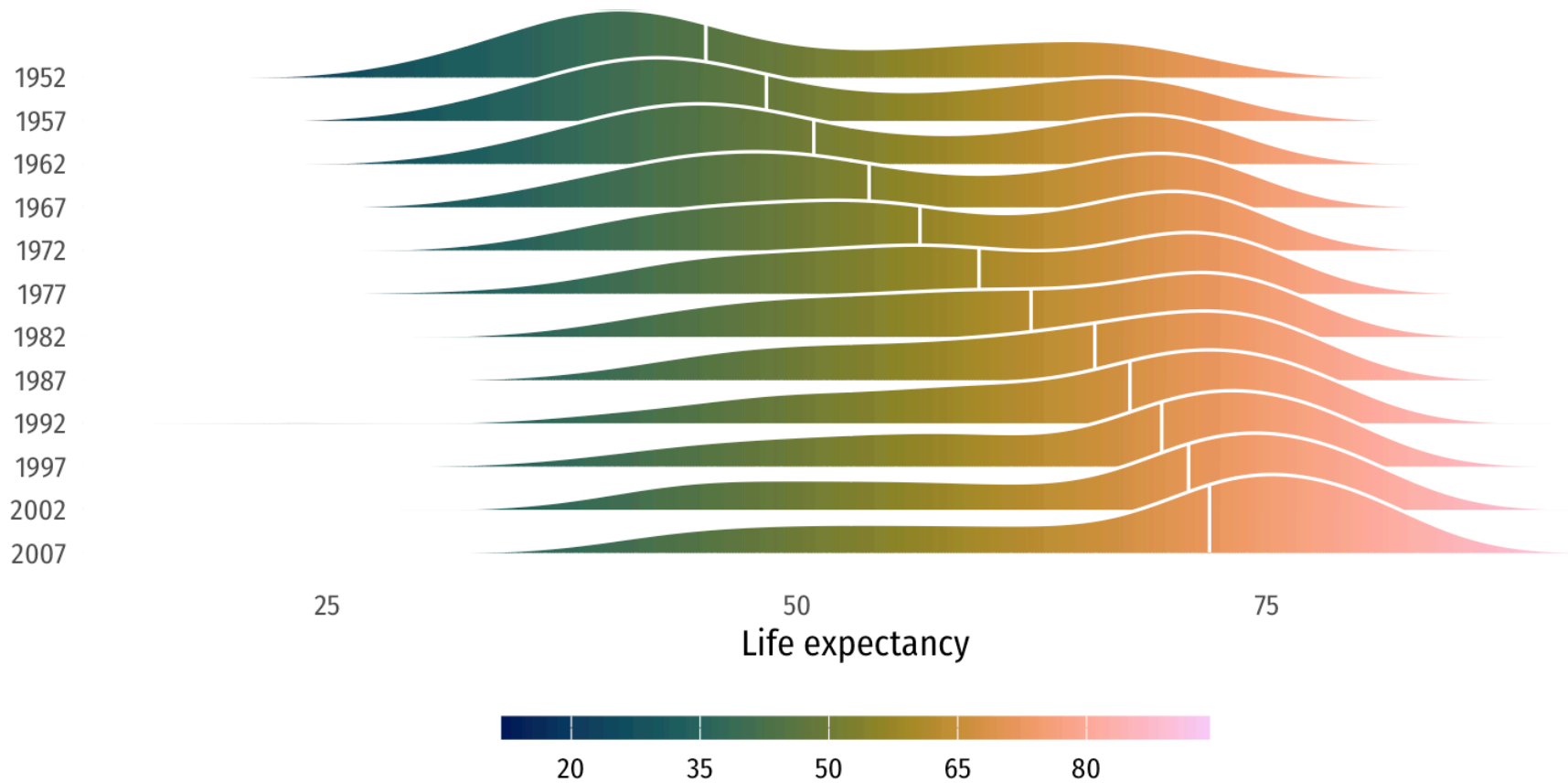


Figure 8.28 in Alberto Cairo's *The Truthful Art*: Heat map by Jorge Camões

Time on y-axis + geom_density()



Time in animation + geom_point()



Time in maps

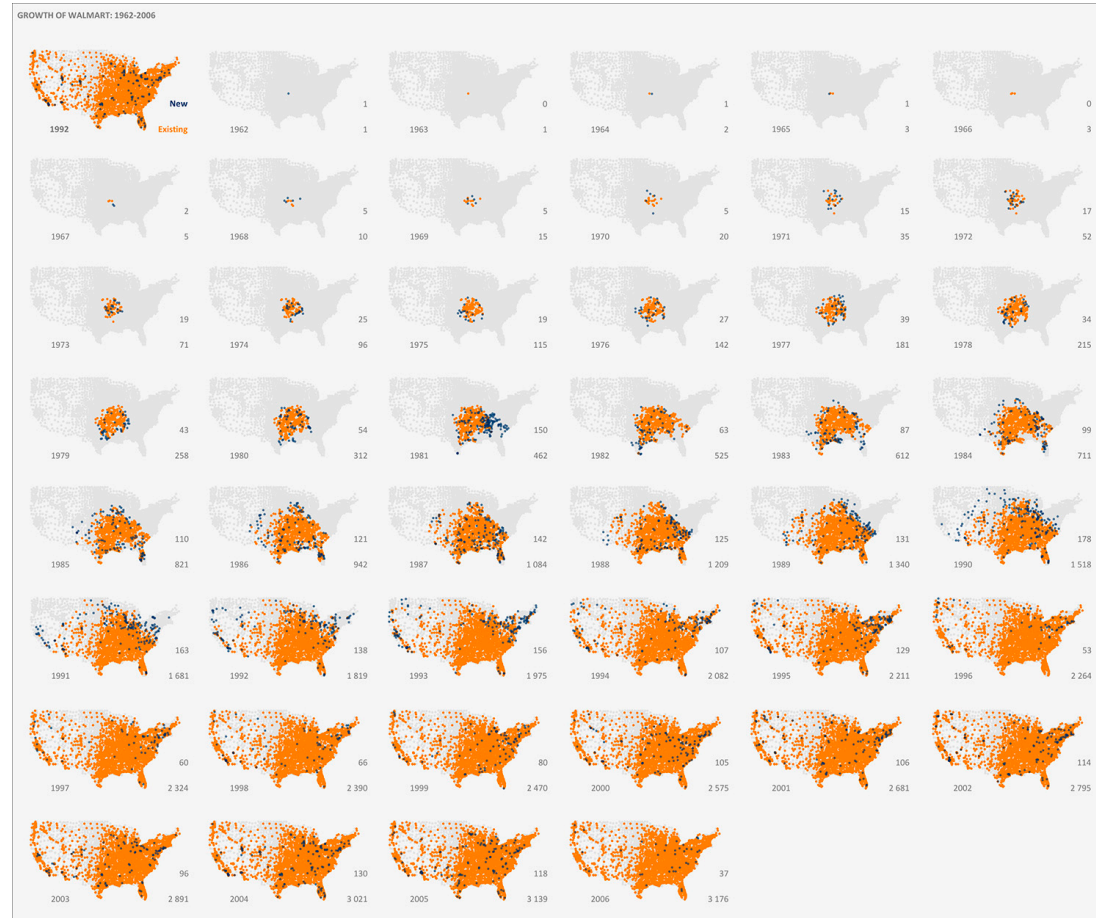
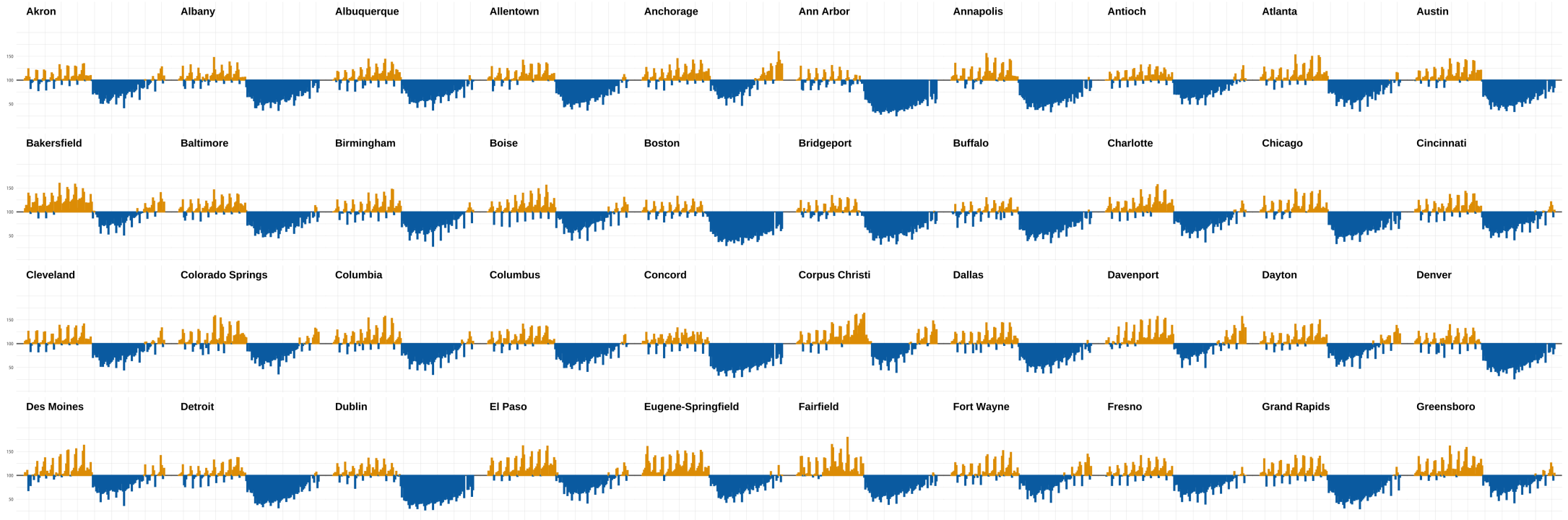


Figure 8.30 in Alberto Cairo's *The Truthful Art: Map of the spread of Walmart* by Jorge Camões

Time in `geom_col()` + small multiples

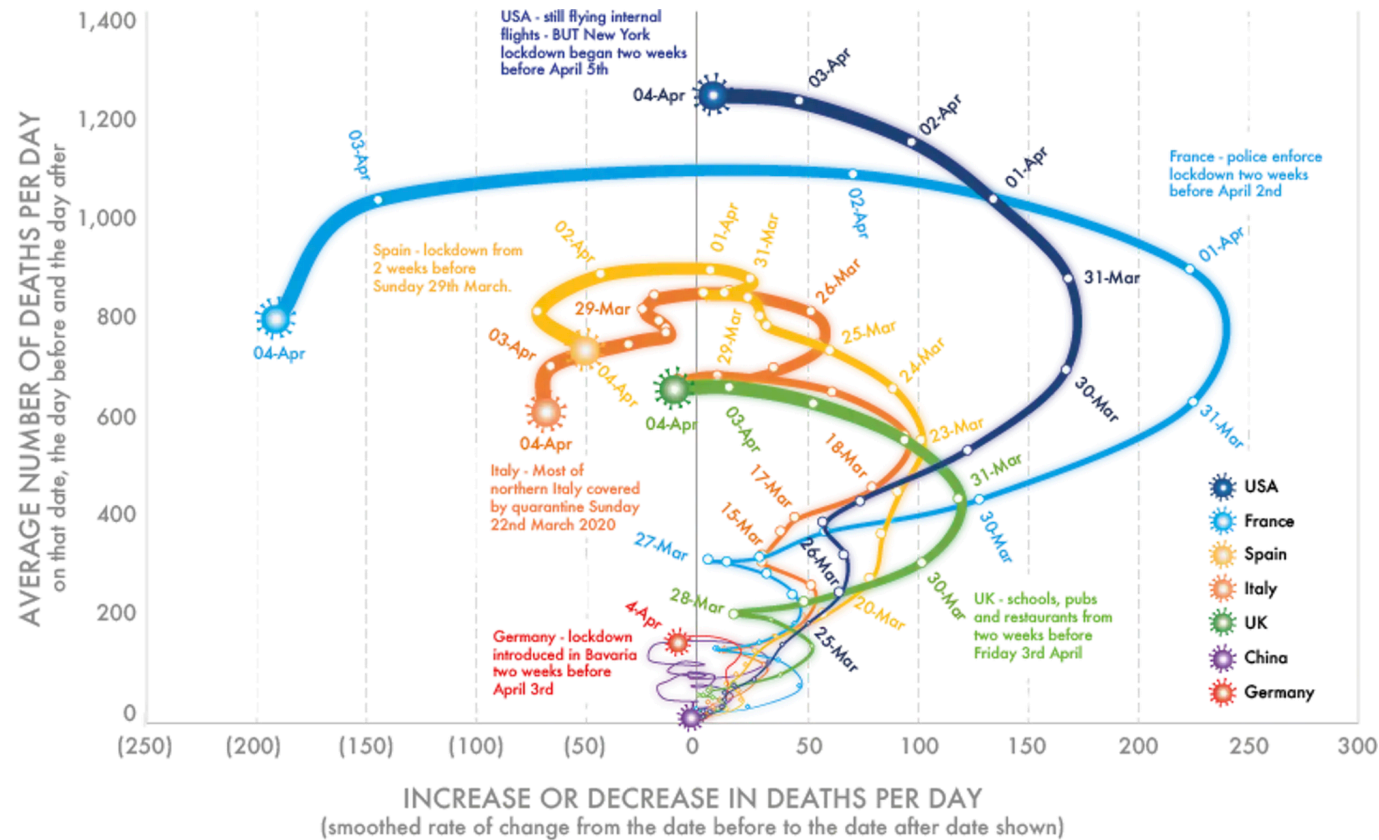
Driving Trends in One Hundred American Cities, January 13th - May 18th, 2020

Data are indexed to 100 for each city's usage on January 13th.



Kieran Healy, "The Kitchen Counter Observatory"

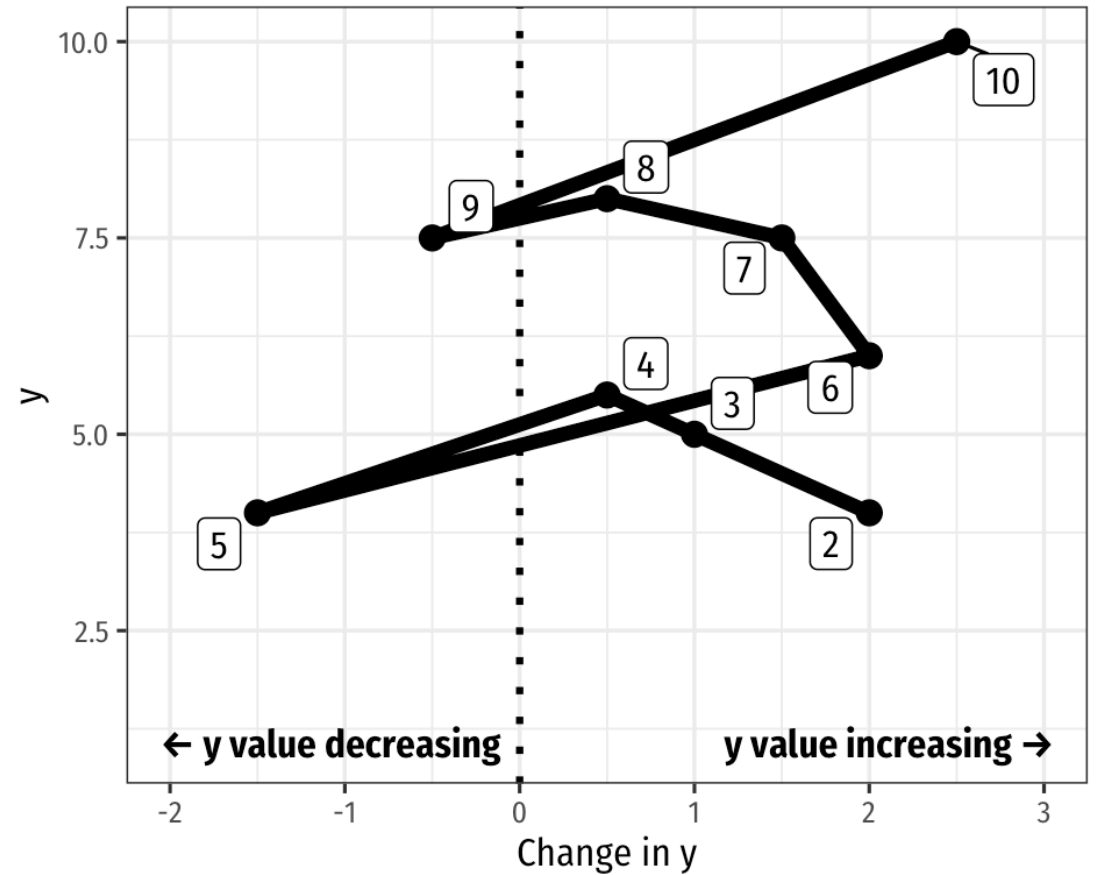
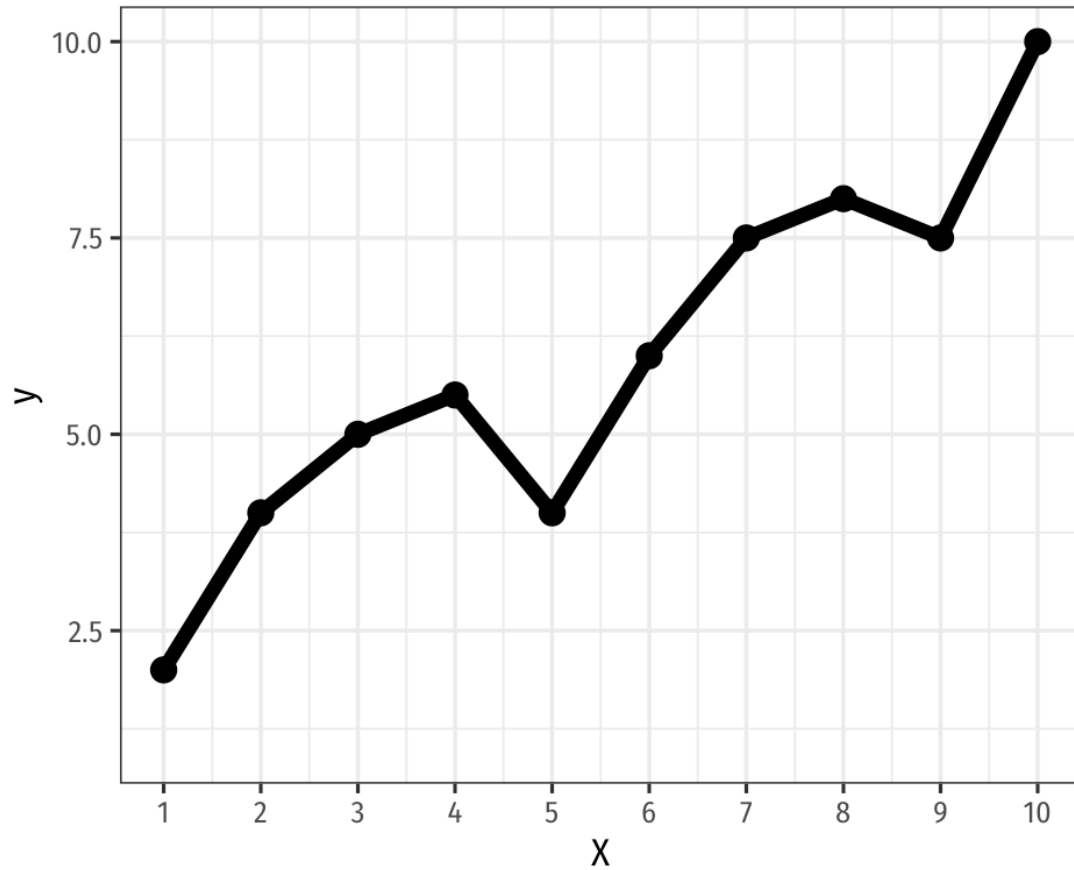
Don't go wild with time mapping!



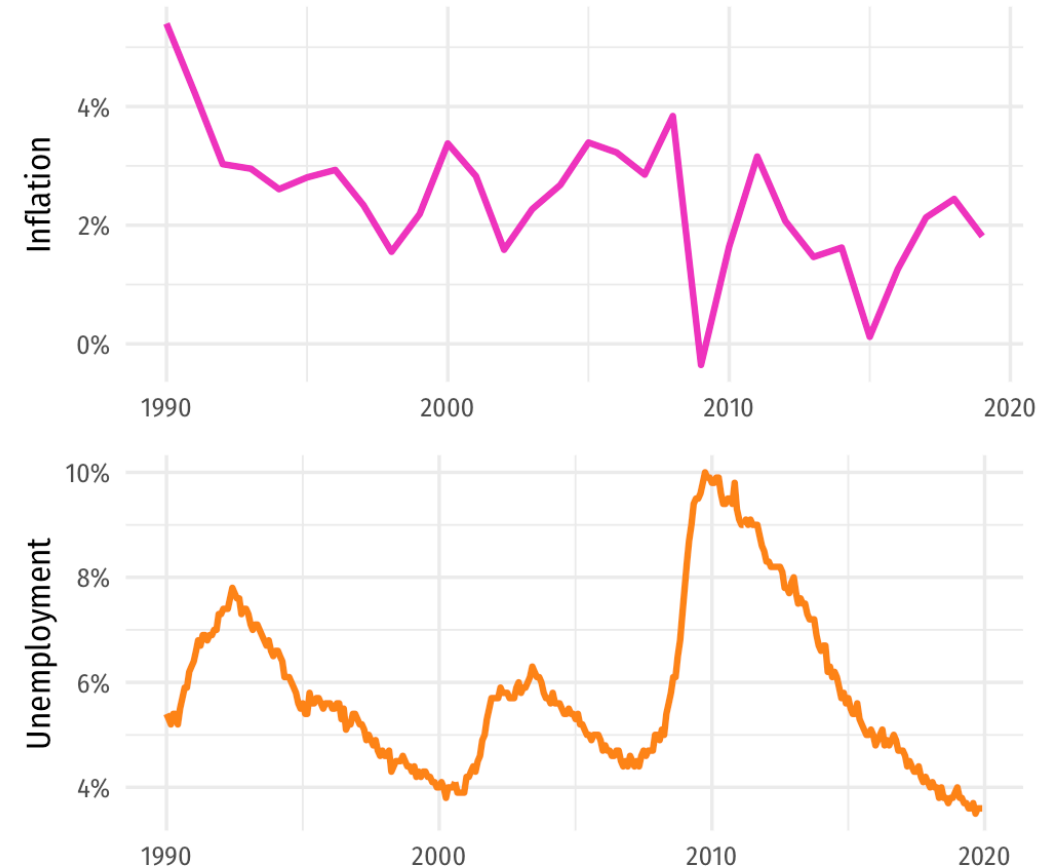
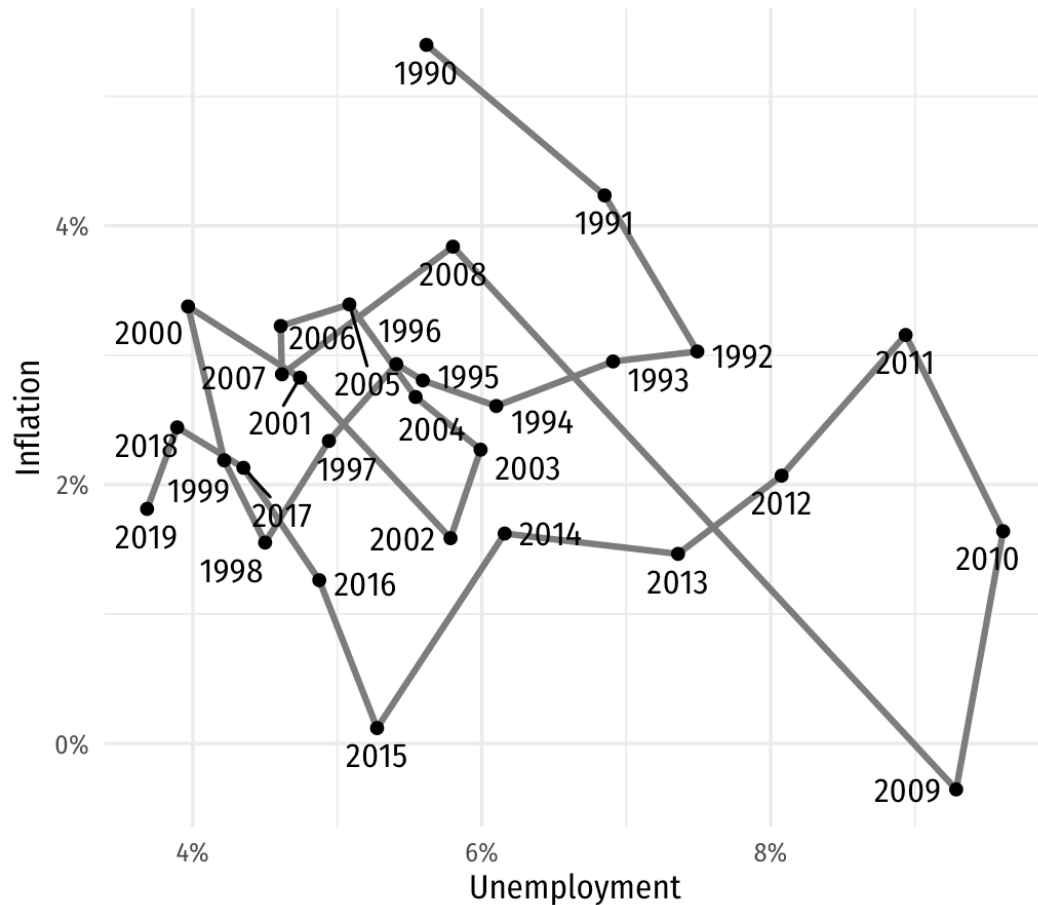
Danny Dorling/Kirsten McClure, Author provided

Tornado plot: When a curve crosses the left of the central axis, the number of deaths per day falls

Interpreting tornado plots



Better with multiple plots



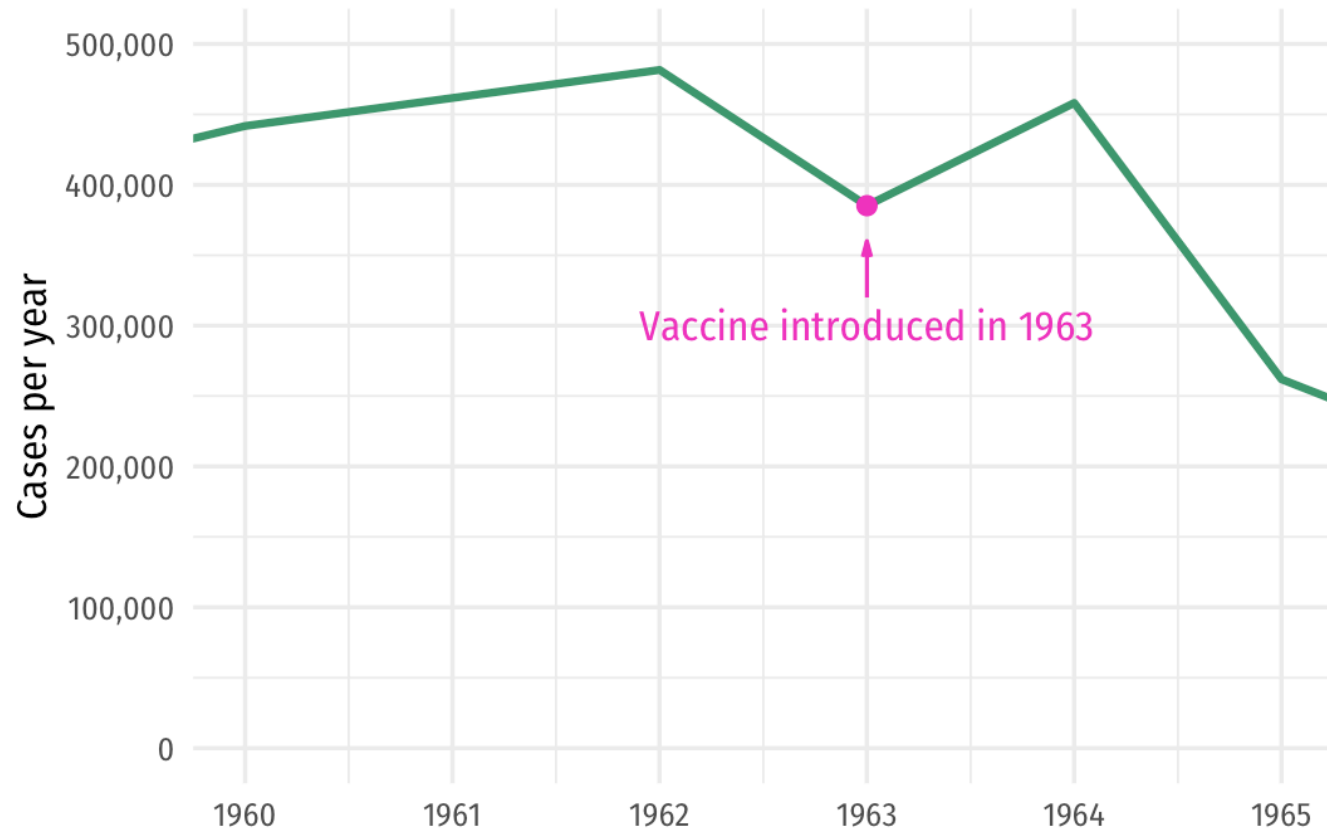
Starting, ending, and decomposing time



**You always have to choose
a start and end point**

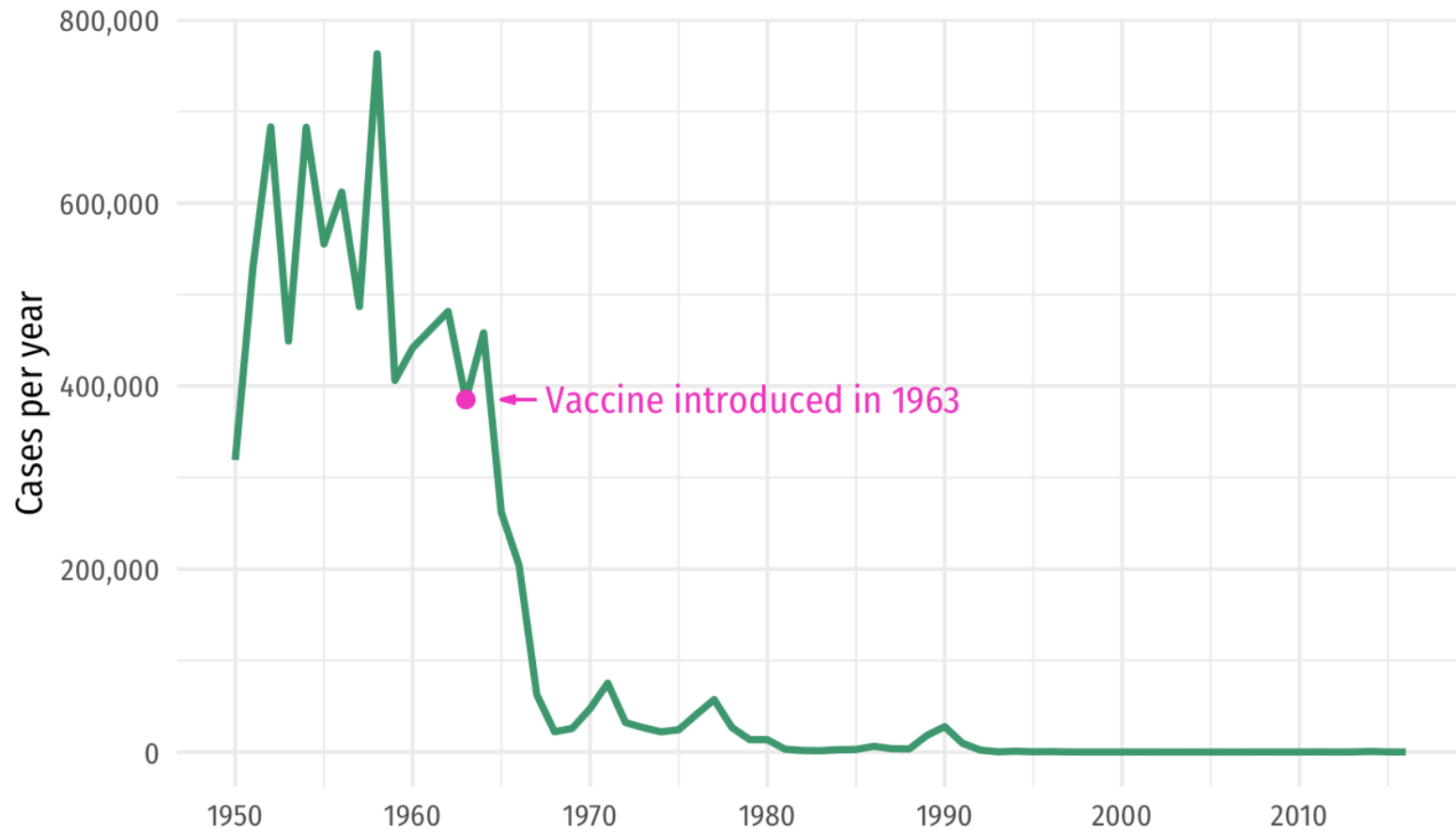
**Start and end at reasonable times that
help maintain the context of the story**

Measles vaccine was pretty effective!



Source: CDC, Epidemiology and Prevention of Vaccine-Preventable Diseases, 13th Edition

Measles vaccine was *incredible!*

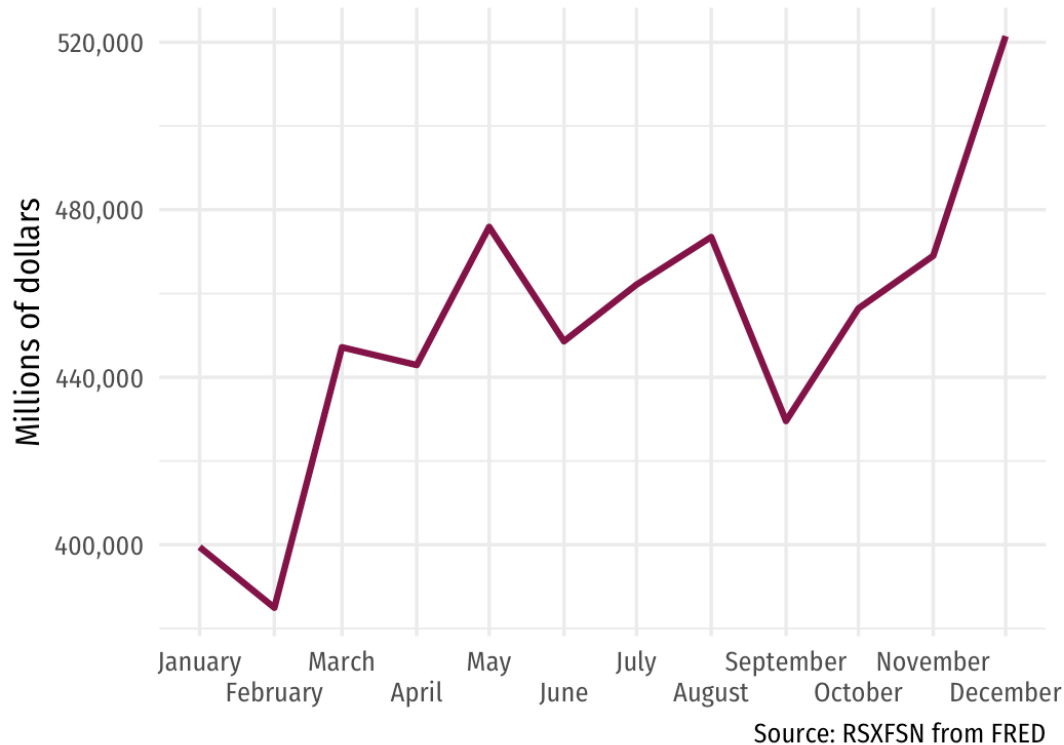


Source: CDC, Epidemiology and Prevention of Vaccine-Preventable Diseases, 13th Edition

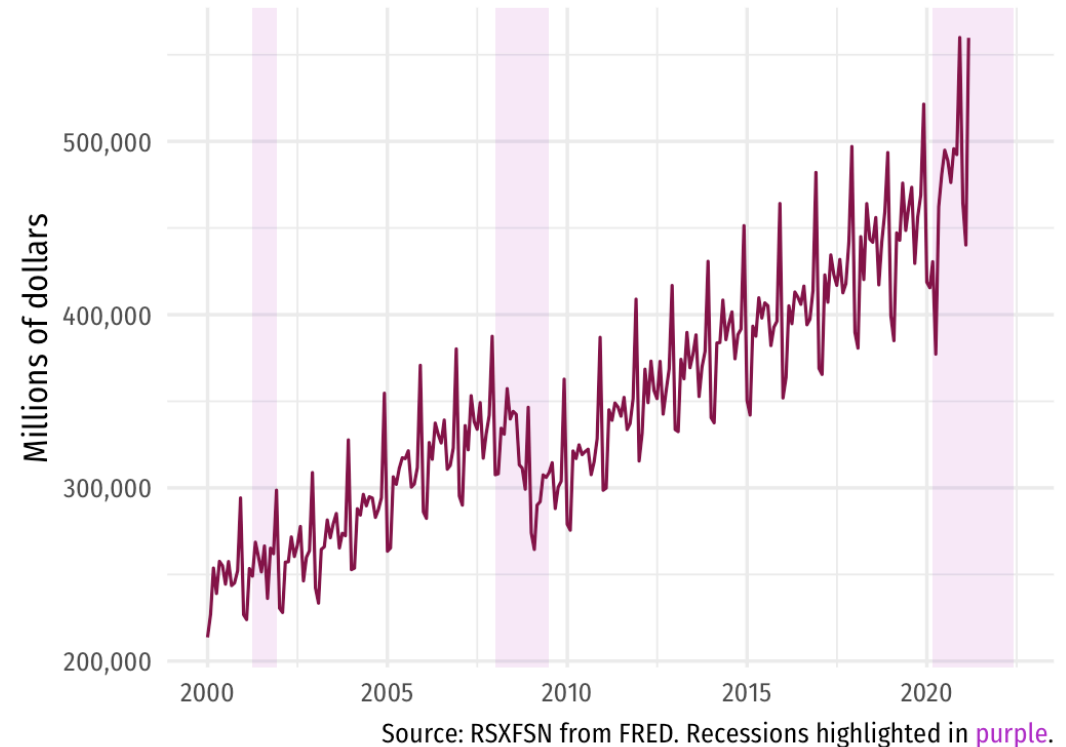
Seasonality

Don't mistake seasonality for actual trends

Total 2019 retail sales in the United States

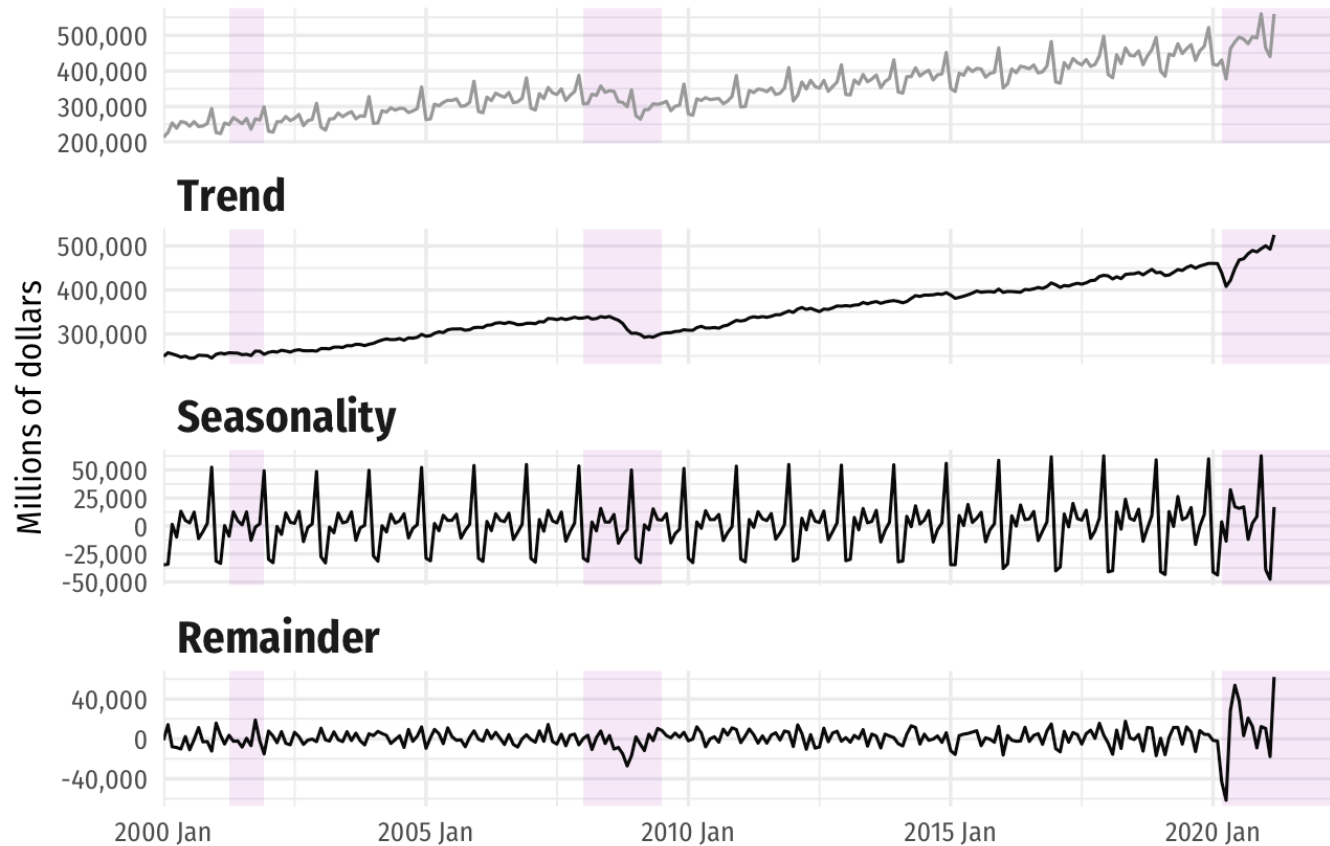


Total retail sales in the United States, 2000–2019



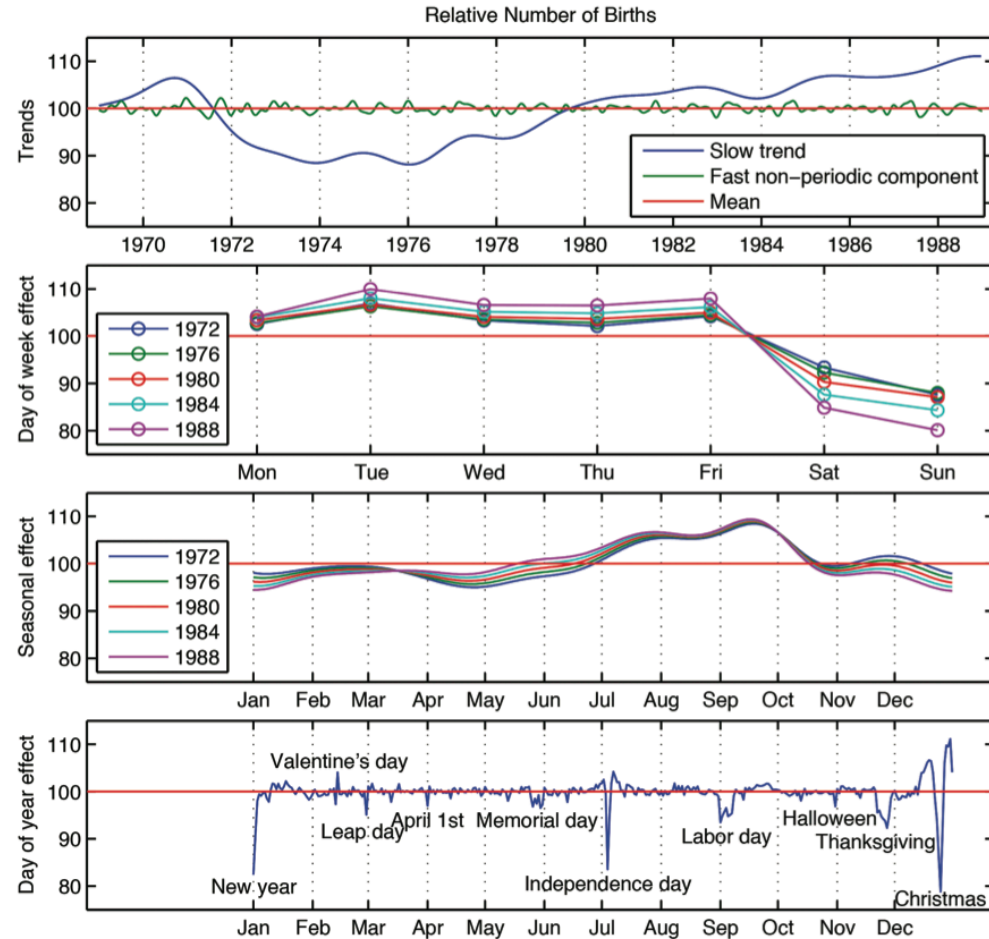
Decomposition

Total retail sales in the United States, 2000–2019



Source: RSXFSN from FRED. Recessions highlighted in purple.

Birthday decomposition



Cover of Andrew Gelman, et al., *Bayesian Data Analysis*