Space

Session 12

PMAP 8921: Data Visualization with R
Andrew Young School of Policy Studies
May 2020

Plan for today
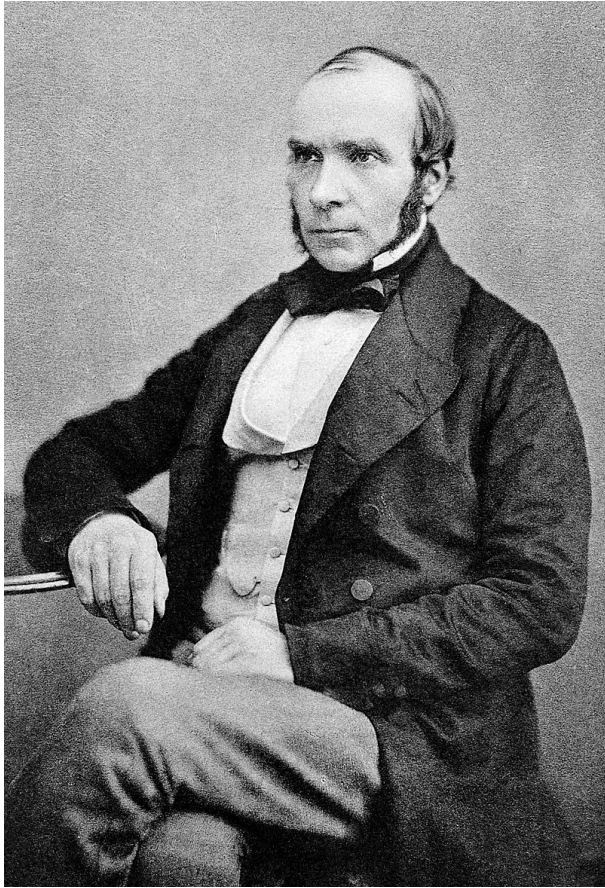
Maps and truth

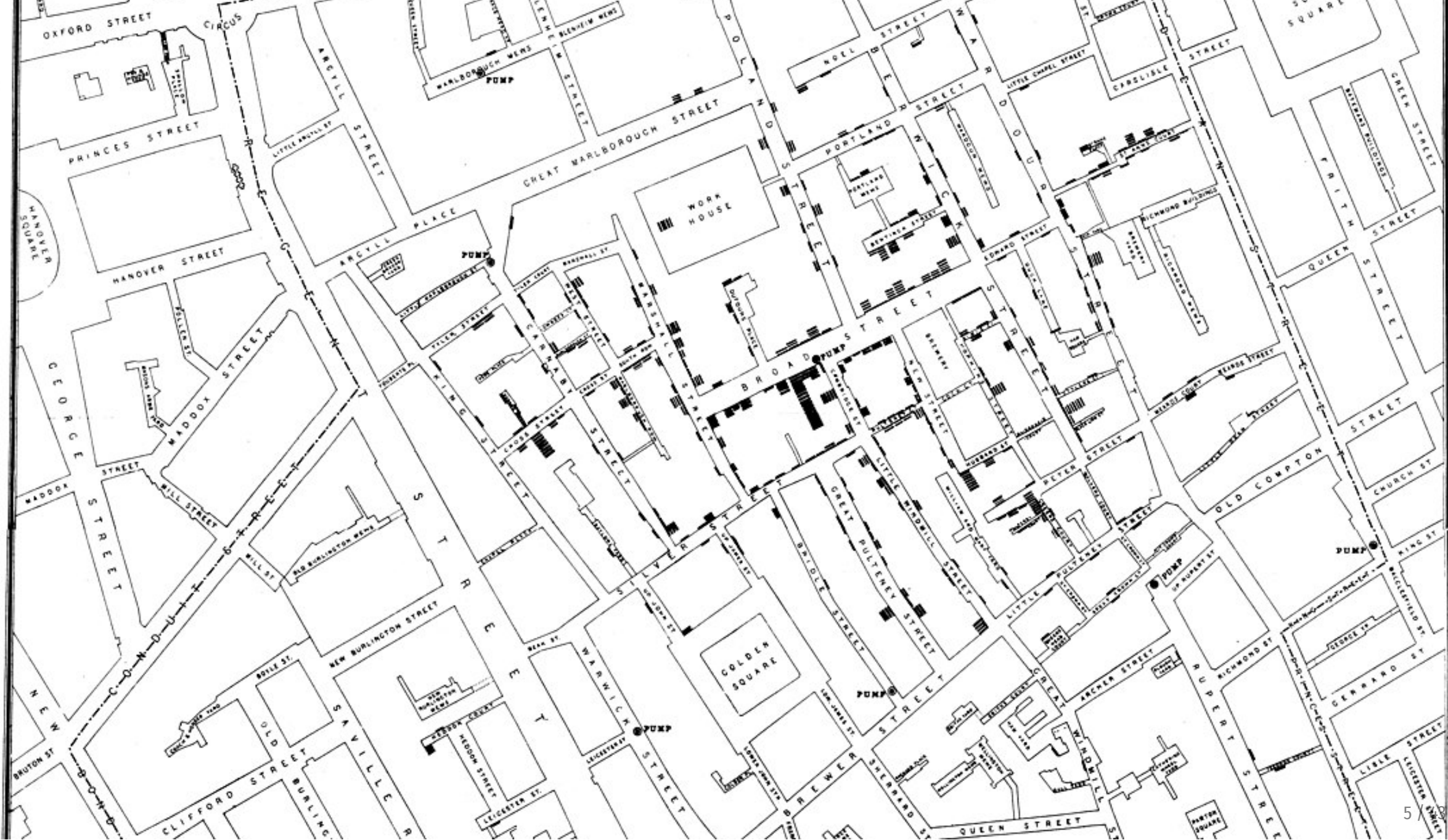Putting data on maps

GIS in R with sf
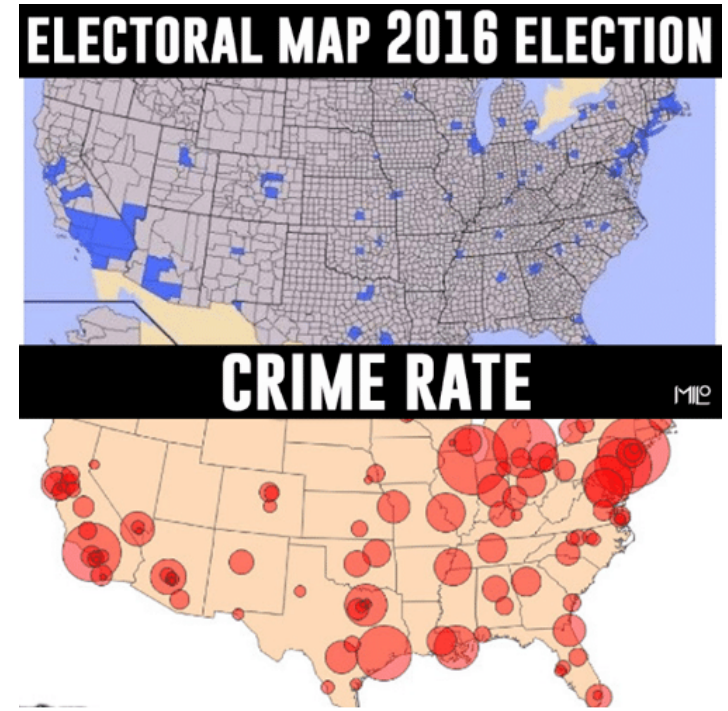
Maps and truth

John Snow and 1854 cholera epidemic

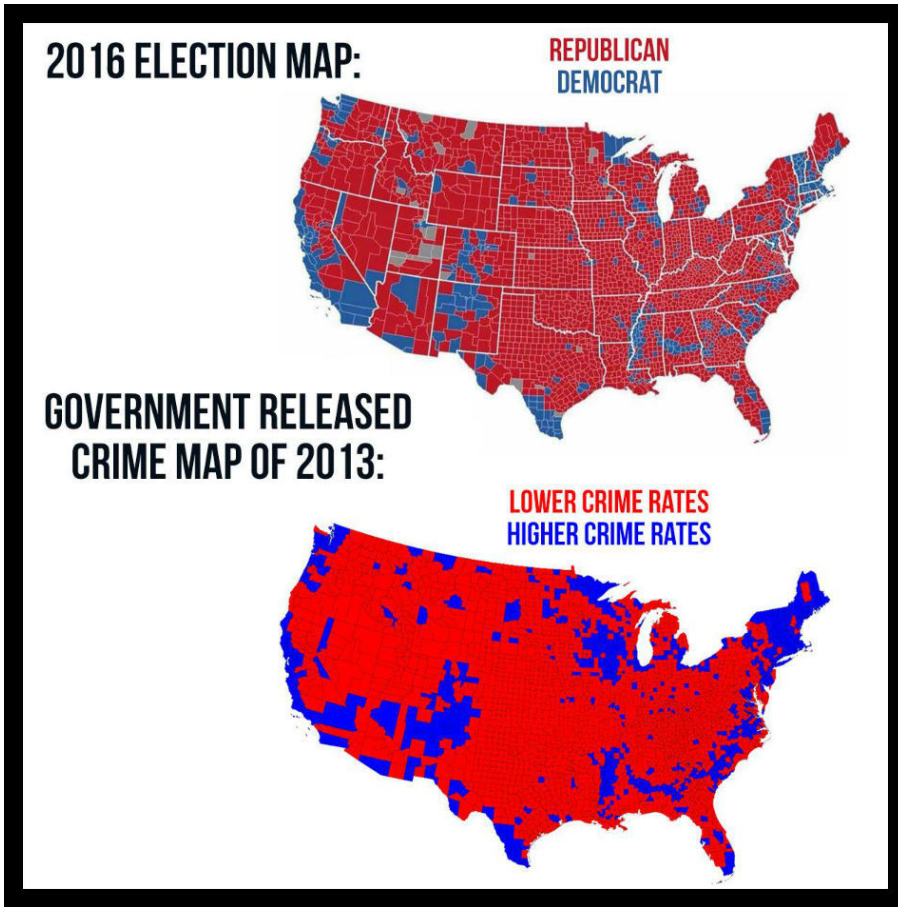
This Jo(h)n Snow knows things

10% of the population of Soho died in a week (!!)

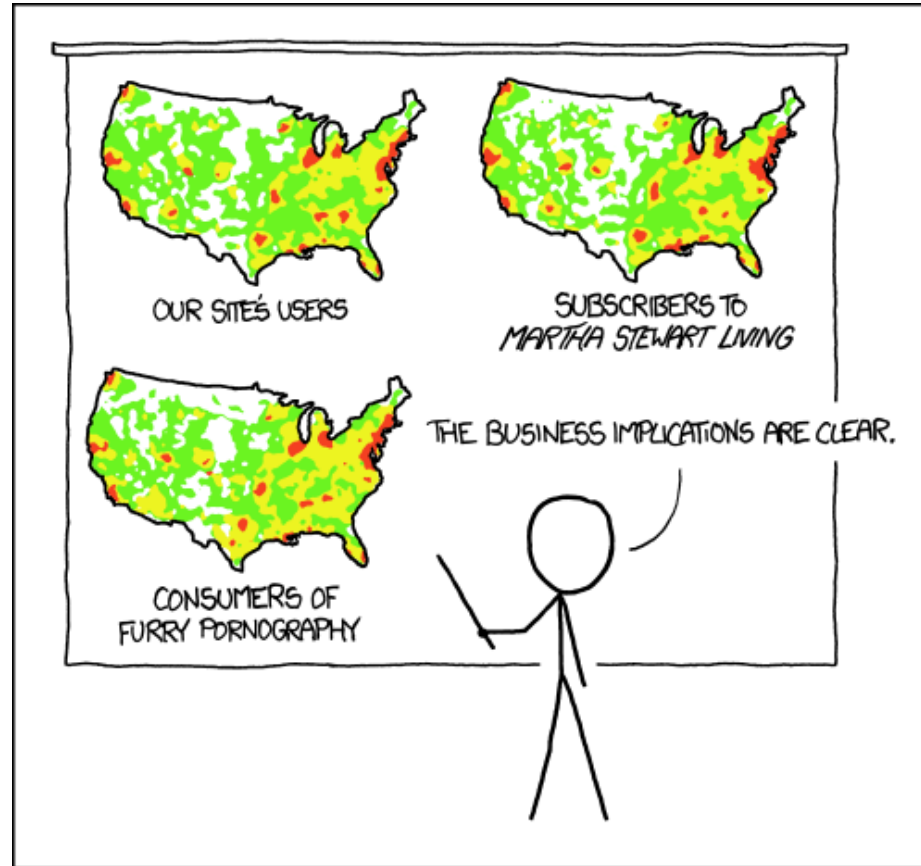**Miasma theory said it was because the air was bad**

OXFORD STREET CIRCUS

SOHO SQUARE

ARGYLL STREET

PRINCES STREET

HANOVER STREET

HANOVER SQUARE

MADDOX STREET

GEORGE STREET

MADDOX STREET

BILL STREET

WILL ST.

OLD BURLINGTON MEWS

NEW BURLINGTON STREET

NEW BURLINGTON MEWS

BOYLE ST.

SAVILLE ROW

CLIFFORD STREET

BURLINGTON STREET

OLD BURLINGTON STREET

BRUTON ST.

NEW CONDUIT STREET

KING STREET

TYLER STREET

FOUBERTS PLACE

CARNABY STREET

ARGYLL PLACE

LITTLE ARGYLL ST.

MARLBOROUGH MEWS

MARLBOROUGH STREET

GREAT MARLBOROUGH STREET

BLENHEIM MEWS

NOEL STREET

POLAND STREET

PORTLAND STREET

WORK HOUSE

SILVER STREET

WARWICK STREET

BRIDLE STREET

GREAT PULTENEY STREET

LITTLE PULTENEY STREET

GOLDEN SQUARE

BROAD STREET

PETER STREET

GREAT WINDMILL STREET

BREWER STREET

SHERRARD STREET

ARCHER STREET

GREAT WINDMILL STREET

QUEEN STREET

OLD COMPTON STREET

CHURCH ST.

QUEEN STREET

FRITH STREET

CREEK STREET

RICHMOND BUILDINGS

CARLISLE STREET

LITTLE CHAPEL STREET

WARDOUR STREET

PORTLAND MEWS

BERWICK STREET

EDWARD STREET

RUPERT STREET

RICHMOND ST.

GERRARD STREET

LISLE STREET

LEICESTER STREET

KING ST.

MACCLESFIELD STREET

GEORGE ST.

PUMP

PUMP

PUMP

PUMP

PUMP

PUMP

PUMP

PUMP

The Broad Street pump

Outright lies

Fake maps and junk maps

"The next great fake news threat? Bot-designed maps"

Points can be useless

Choropleths can be great

Smoky Mountains 2019 Fall Foliage Prediction Map

Choropleths can distort

Land doesn't vote
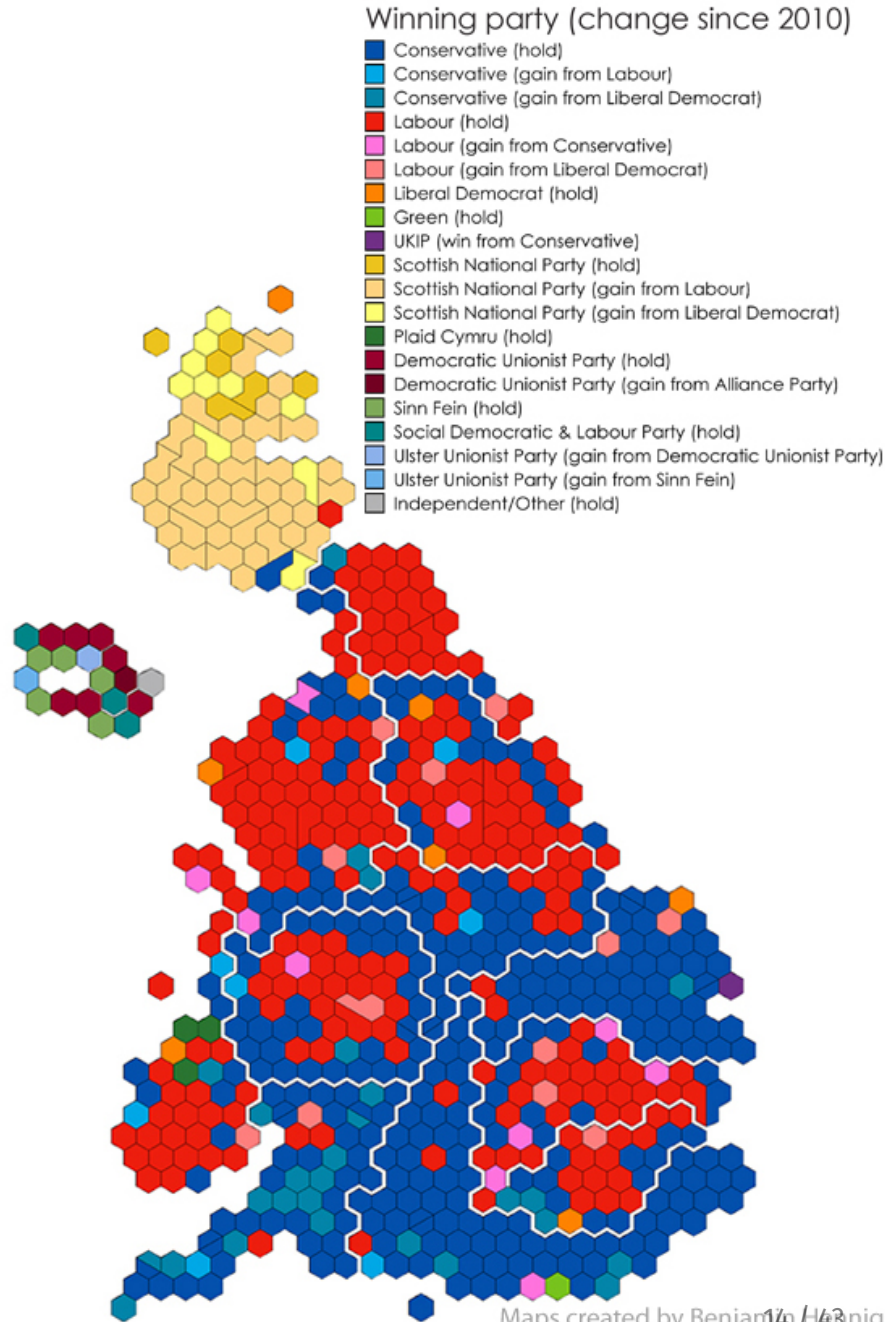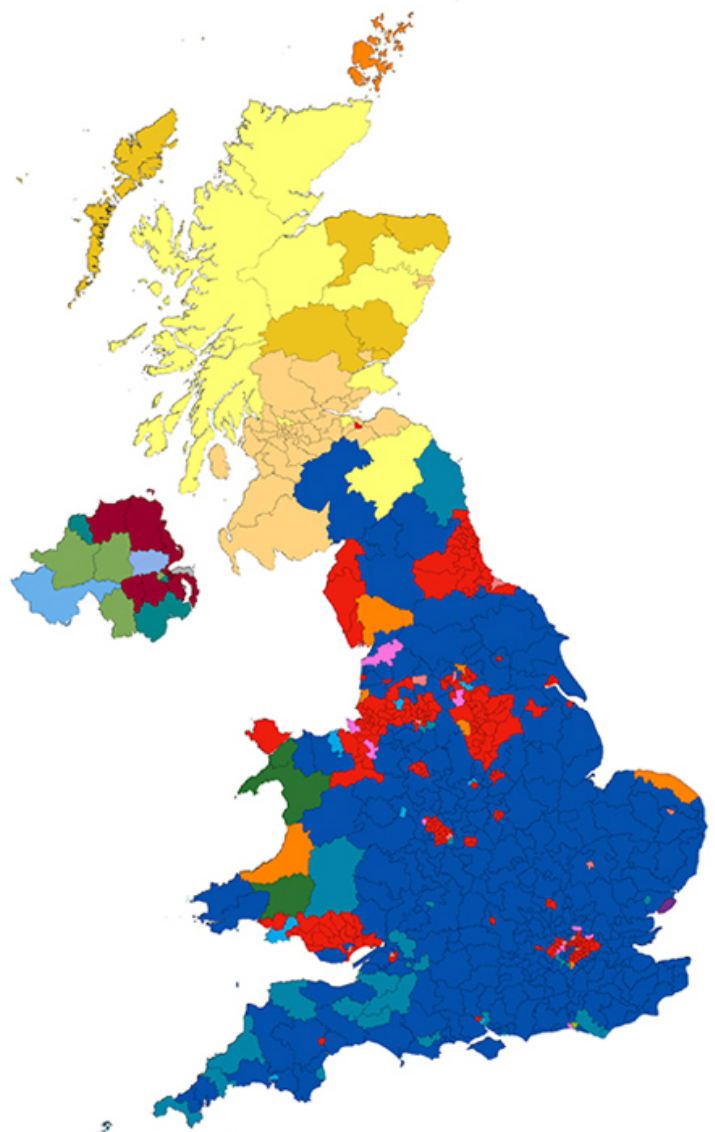

0:00 / 0:07
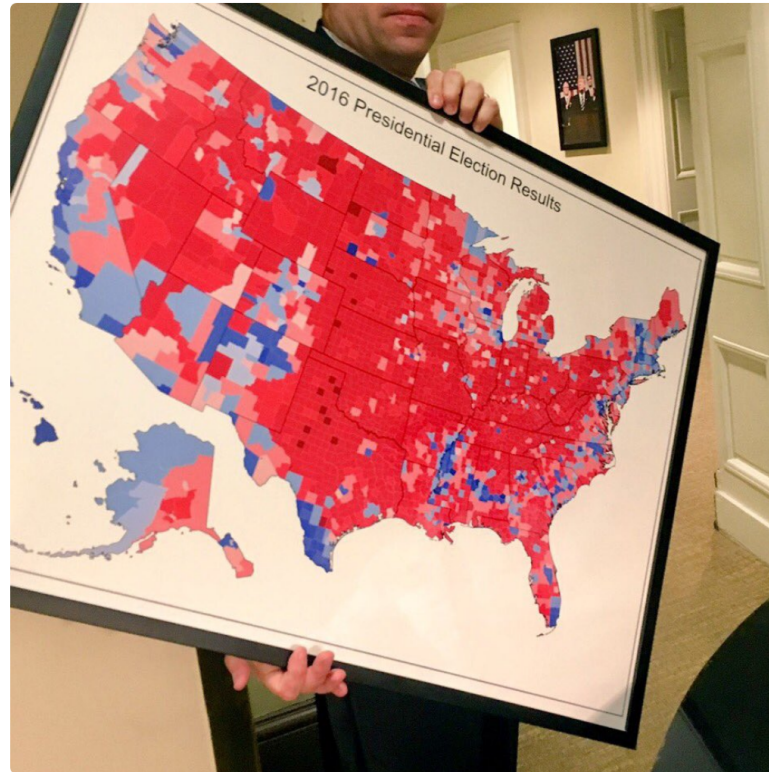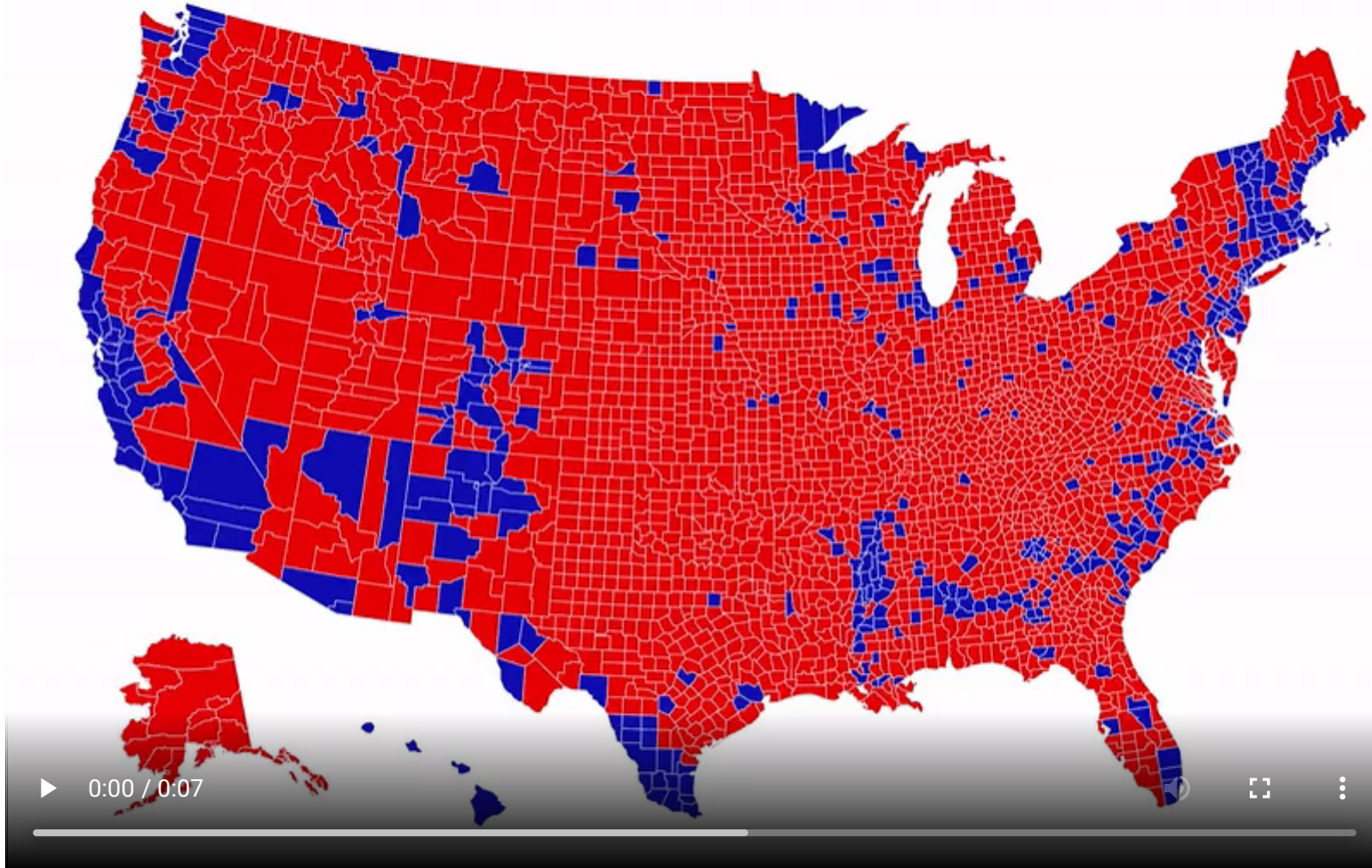
Cartograms
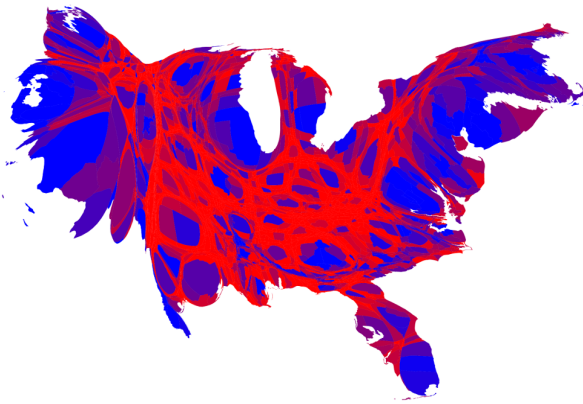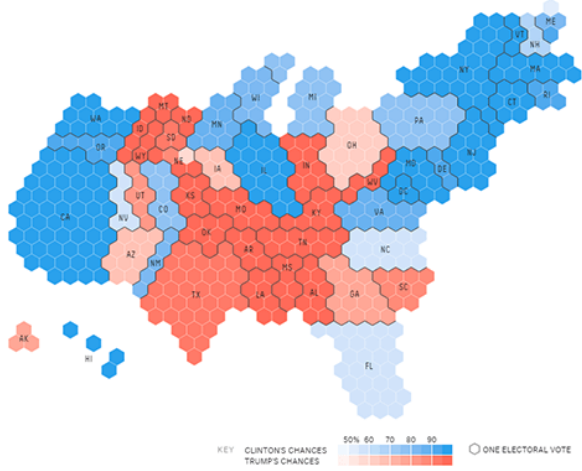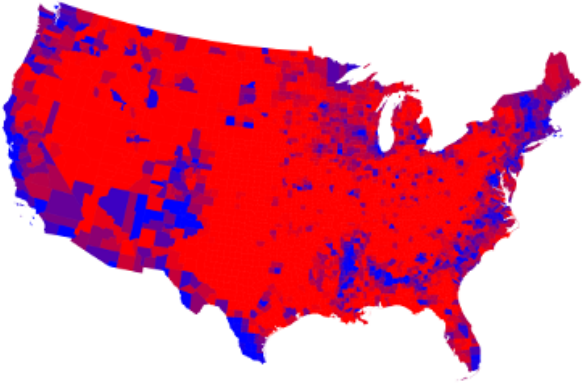
Electoral Doctrine
Mapping the
2015 UK General Election

Winning party (change since 2010)

- ■ Conservative (hold)
- ■ Conservative (gain from Labour)
- ■ Conservative (gain from Liberal Democrat)
- ■ Labour (hold)
- ■ Labour (gain from Conservative)
- ■ Labour (gain from Liberal Democrat)
- ■ Liberal Democrat (hold)
- ■ Green (hold)
- ■ UKIP (win from Conservative)
- ■ Scottish National Party (hold)
- ■ Scottish National Party (gain from Labour)
- ■ Scottish National Party (gain from Liberal Democrat)
- ■ Plaid Cymru (hold)
- ■ Democratic Unionist Party (hold)
- ■ Democratic Unionist Party (gain from Alliance Party)
- ■ Sinn Fein (hold)
- ■ Social Democratic & Labour Party (hold)
- ■ Ulster Unionist Party (gain from Democratic Unionist Party)
- ■ Ulster Unionist Party (gain from Sinn Fein)
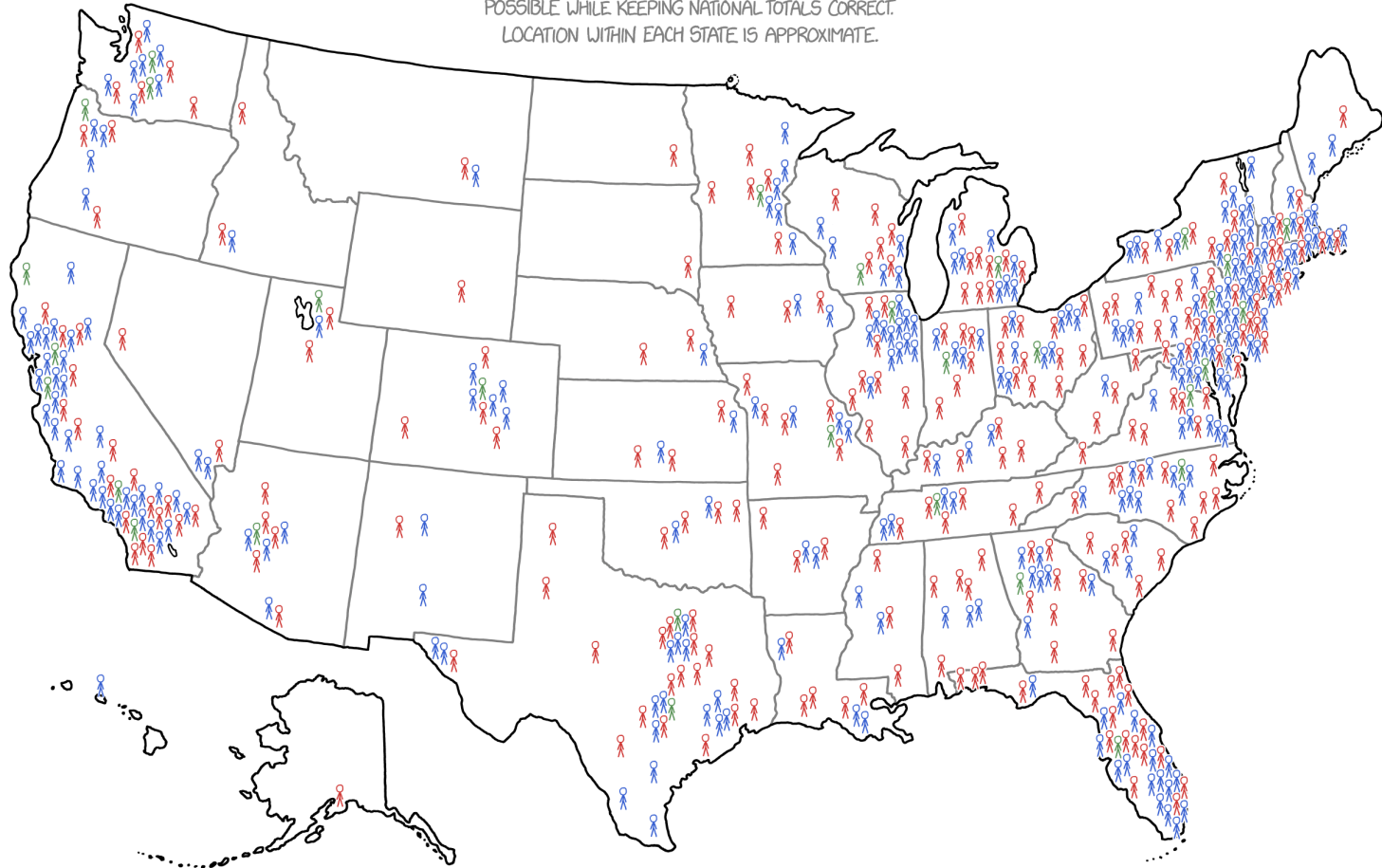- ■ Independent/Other (hold)

2016 ELECTION MAP
EACH FIGURE REPRESENTS 250,000 VOTES
TRUMP   CLINTON   OTHER

VOTES ARE DISTRIBUTED BY STATE AS ACCURATELY AS
POSSIBLE WHILE KEEPING NATIONAL TOTALS CORRECT.
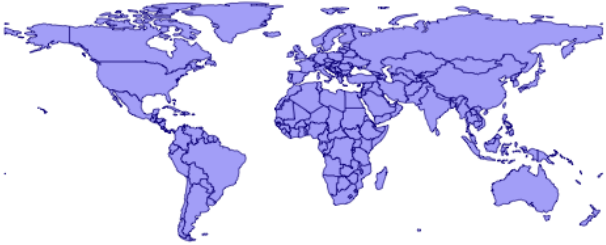
LOCATION WITHIN EACH STATE IS APPROXIMATE.

Projections

Animated world projections

World projections

Longitude-latitude
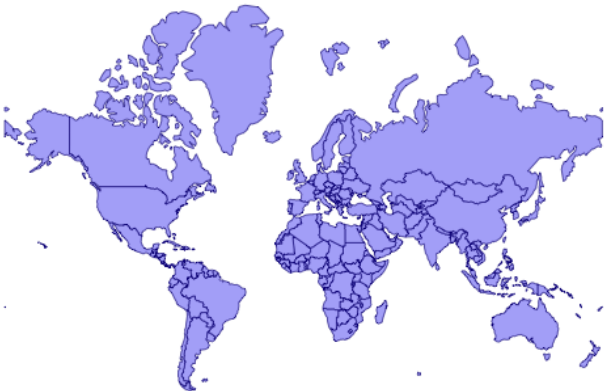


```
crs = "+proj=longlat +ellps=WGS84"
```

Mercator


```
crs = "+proj=merc"
```

Gall-Peters
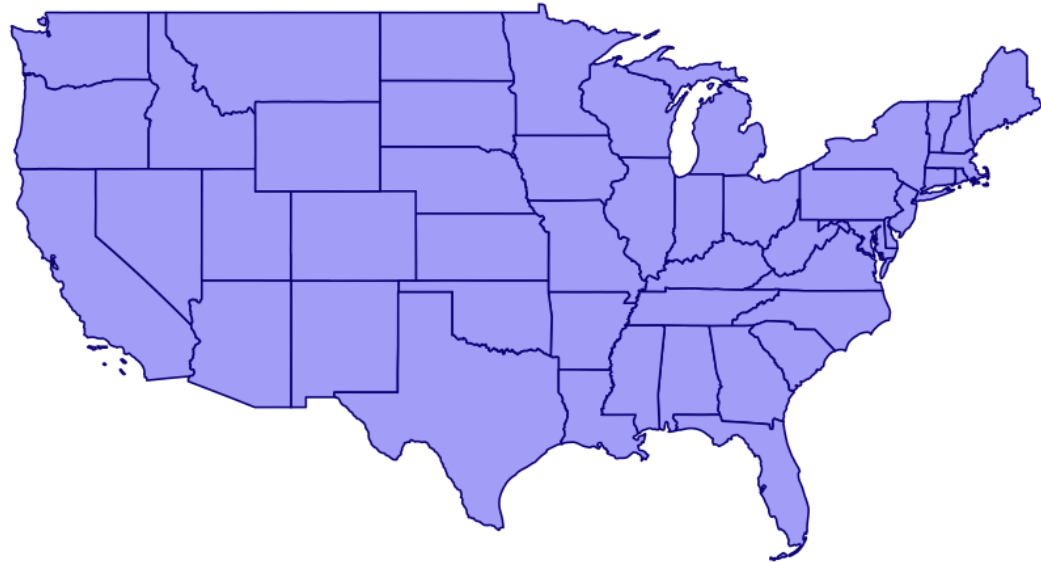


```
crs = 54002
```

Robinson


```
crs = "+proj=robin"
```

US projections

NAD83

crs = 4269

Albers
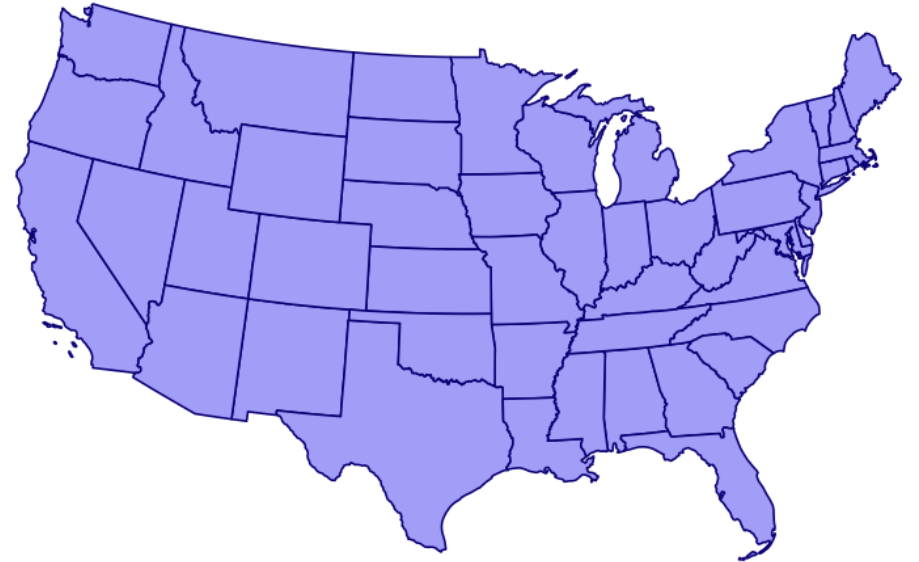
crs = 102003

Finding projection codes

spatialreference.org

epsg.io

proj.org

Most common ones listed on the course website example page

This is an excellent overview of how this all works

And this is a really really helpful overview of all these moving parts

Which projection is best?

None of them

There are no good or bad projections

There are appropriate and inappropriate projections

(but also ew mercator)

Putting data on maps

Maps with lines
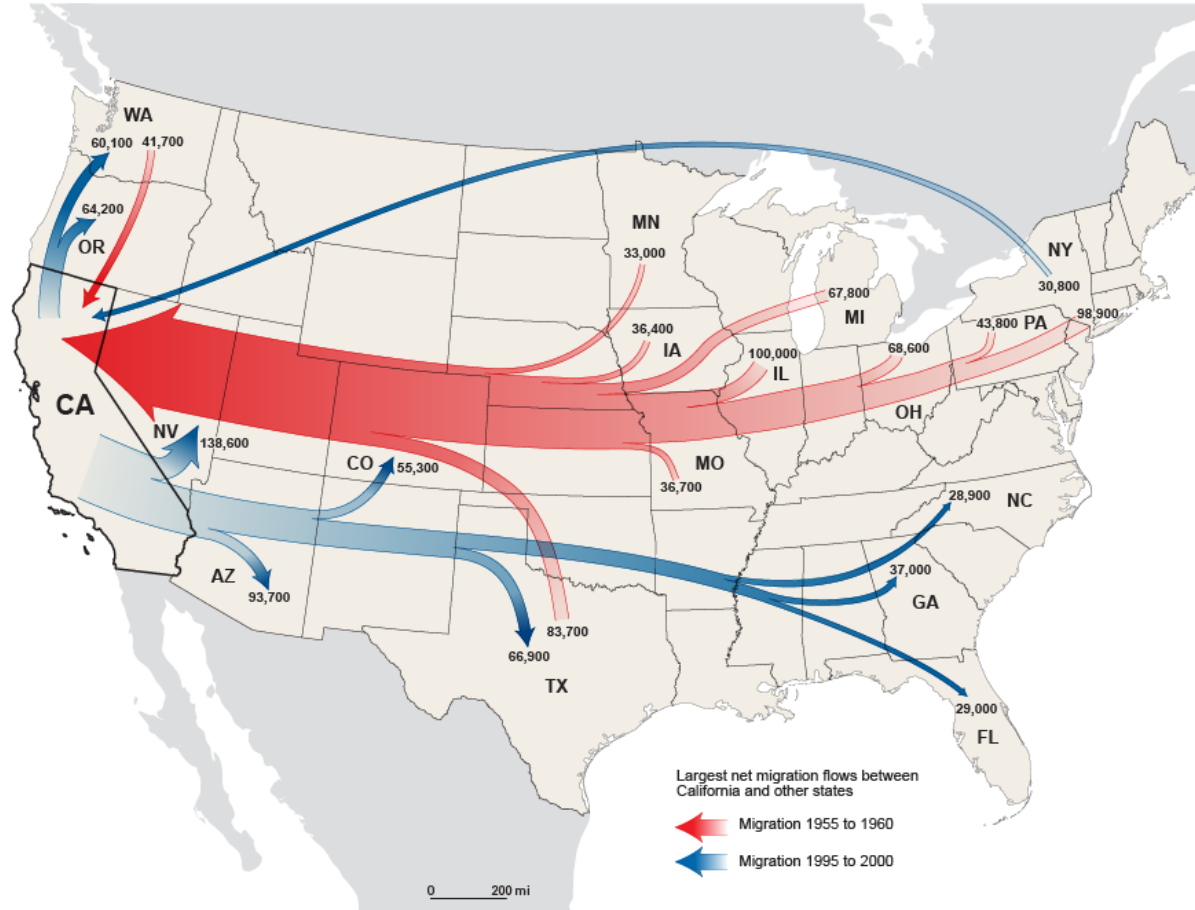
US Census Bureau: Net migration between California and other states

Maps with lines

October 30, 2012
6:59 am EST
(time of forecast download)

top speed: **39.7 mph**
average: **8.4 mph**

1 mph

3 mph

5 mph

10 mph

15 mph

30 mph

Seattle

New York

Chicago

Columbus

Denver

Los Angeles

San Diego

Phoenix

Dallas

San Antonio

Houston

hint.fm Live Wind Map
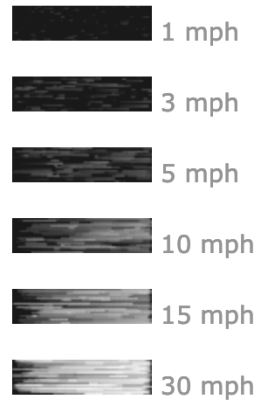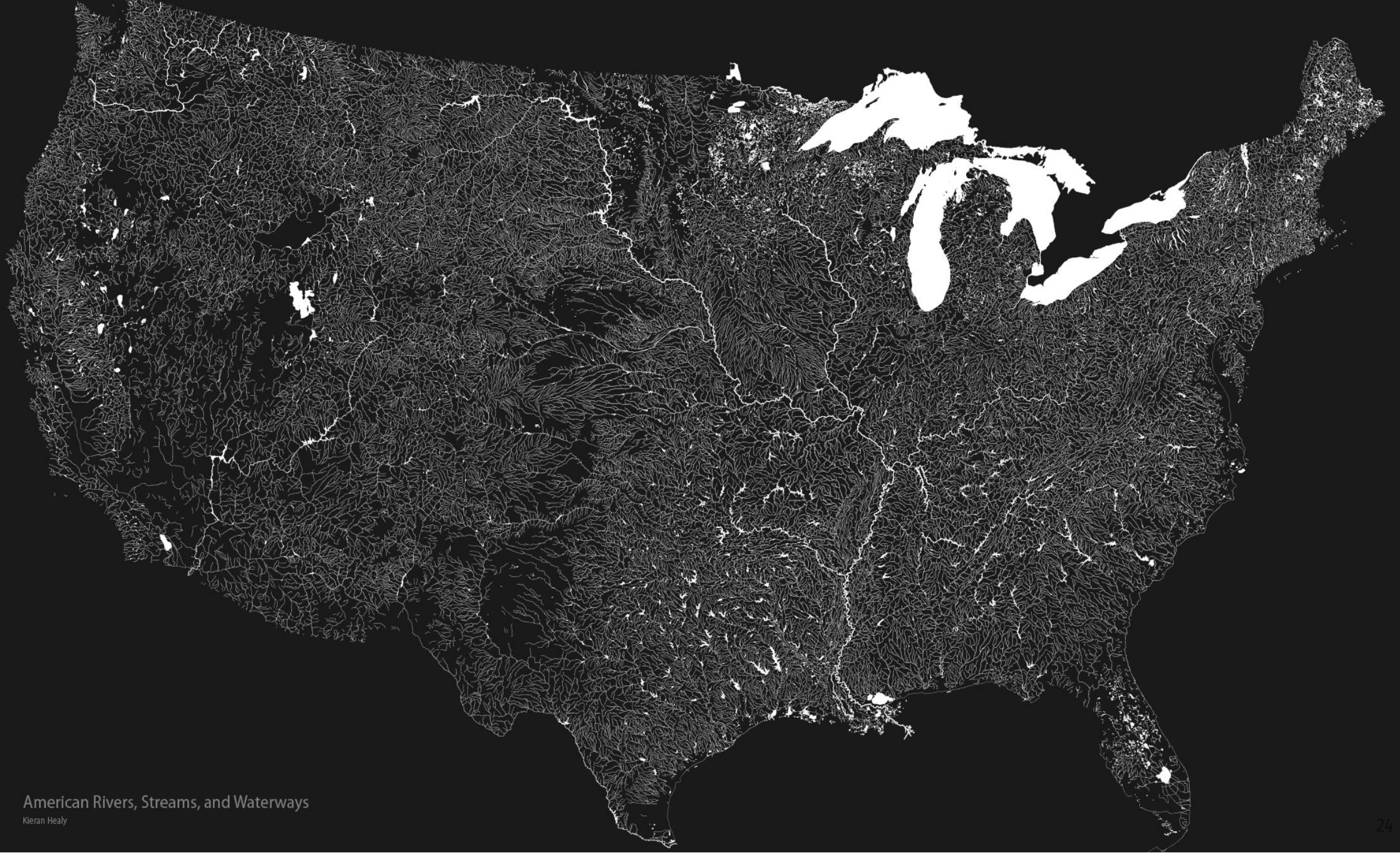
American Rivers, Streams, and Waterways
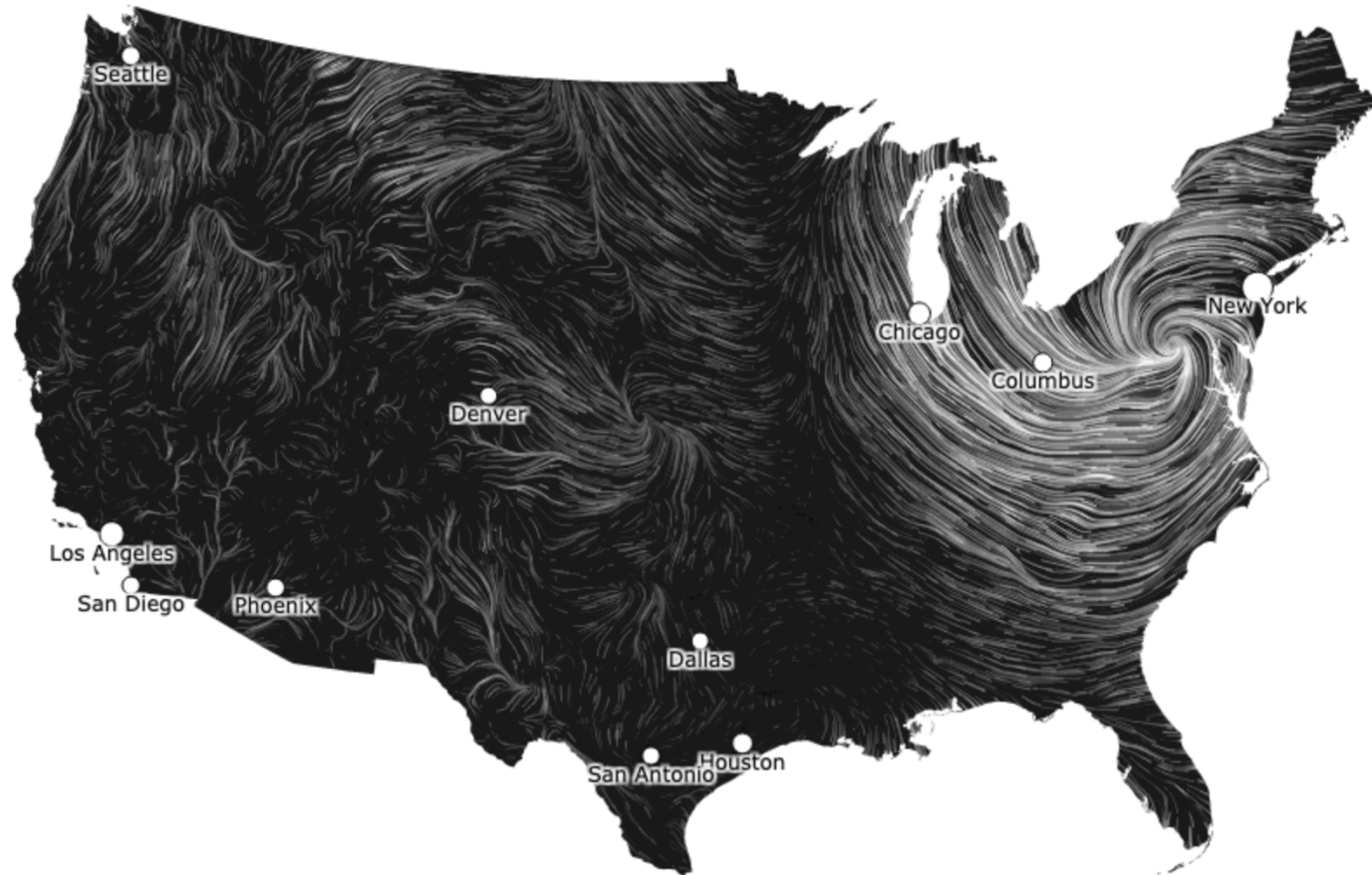Kieran Healy

Maps with points

Every hurricane since 1851, by IDV solutions

Maps with points

Show measles-vaccination rates | Show personal-belief exemptions

60%   80%   92%

Los Angeles

Vaccination rates are low in wealthy, predominantly white neighborhoods, as well as in neighboring Hispanic towns. At Westside Waldorf School, a private school in Pacific Palisades, 16 of the 20 kindergarteners have personal-belief exemptions. In many other schools, the low vaccination rates are not necessarily caused by parents' opting out; children can be conditionally admitted while they catch up on immunizations.
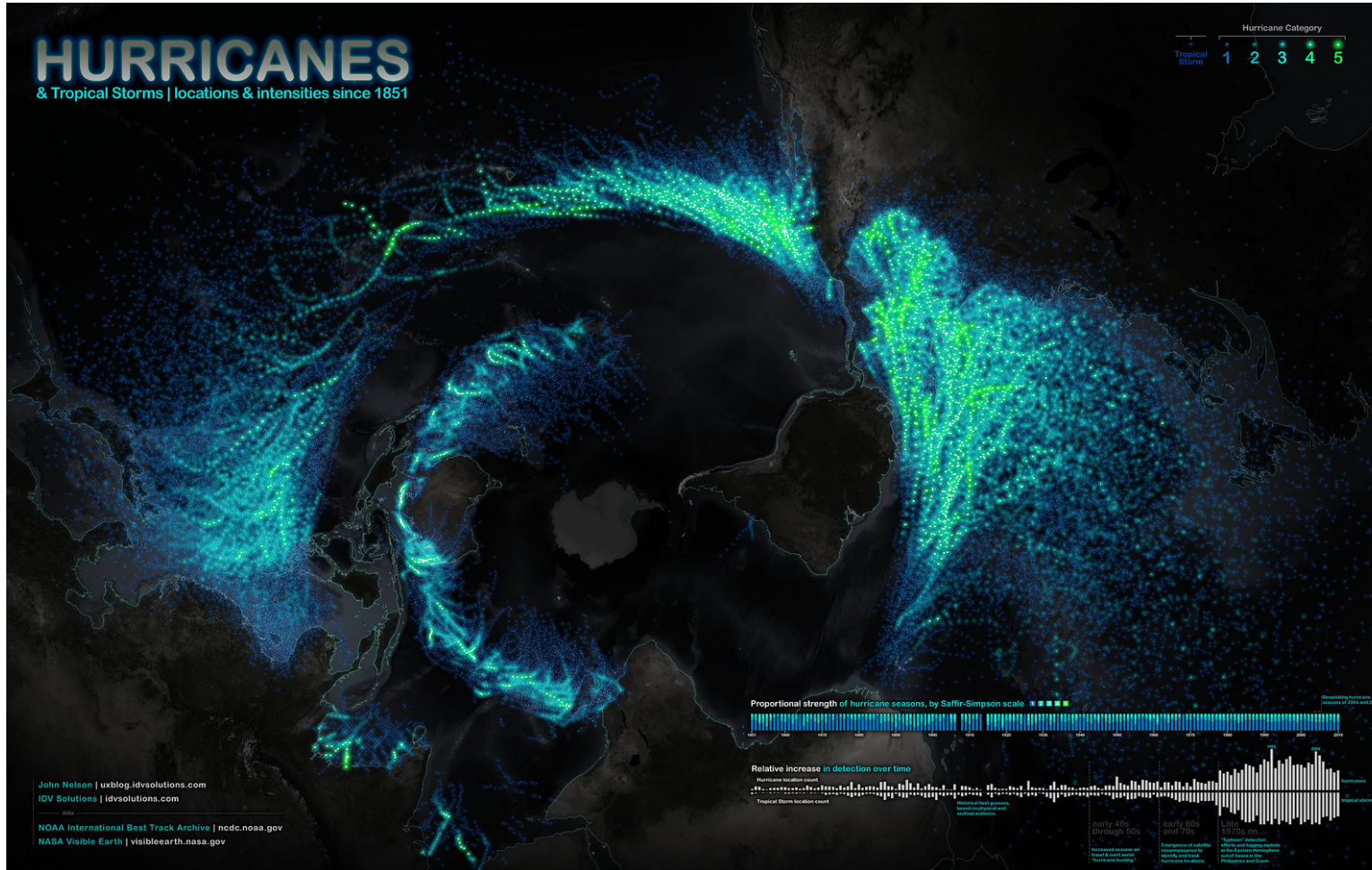
Orange County

The mostly white and affluent towns in South Orange County have low vaccination rates. Some parents have said in surveys that they don't perceive the diseases to be a risk in their community, according to David Nunez, the county medical director for family health.
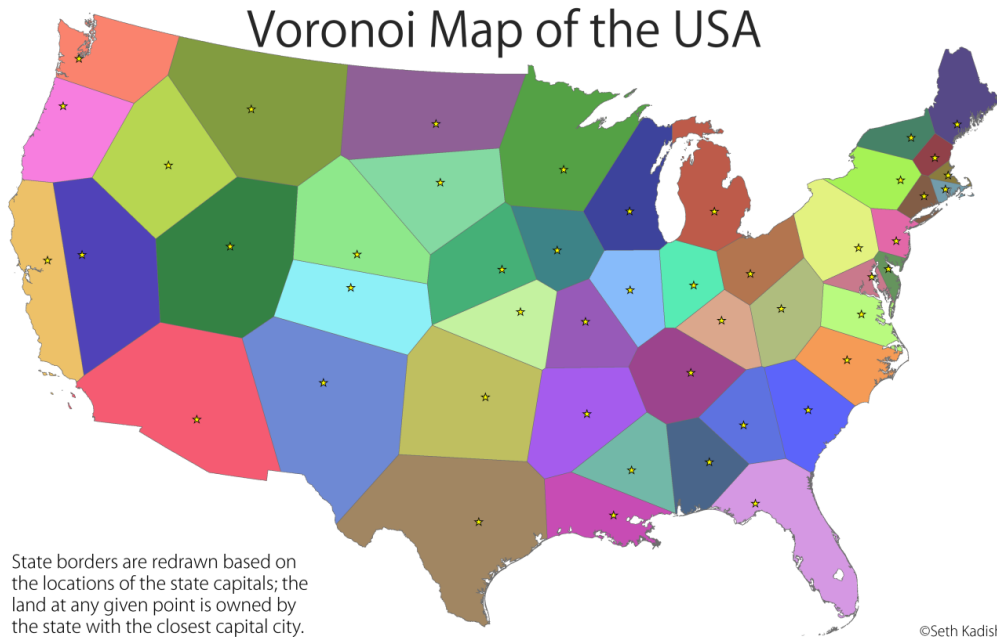
The New York Times, "Vaccination Rates for Every Kindergarten in California

Maps with points
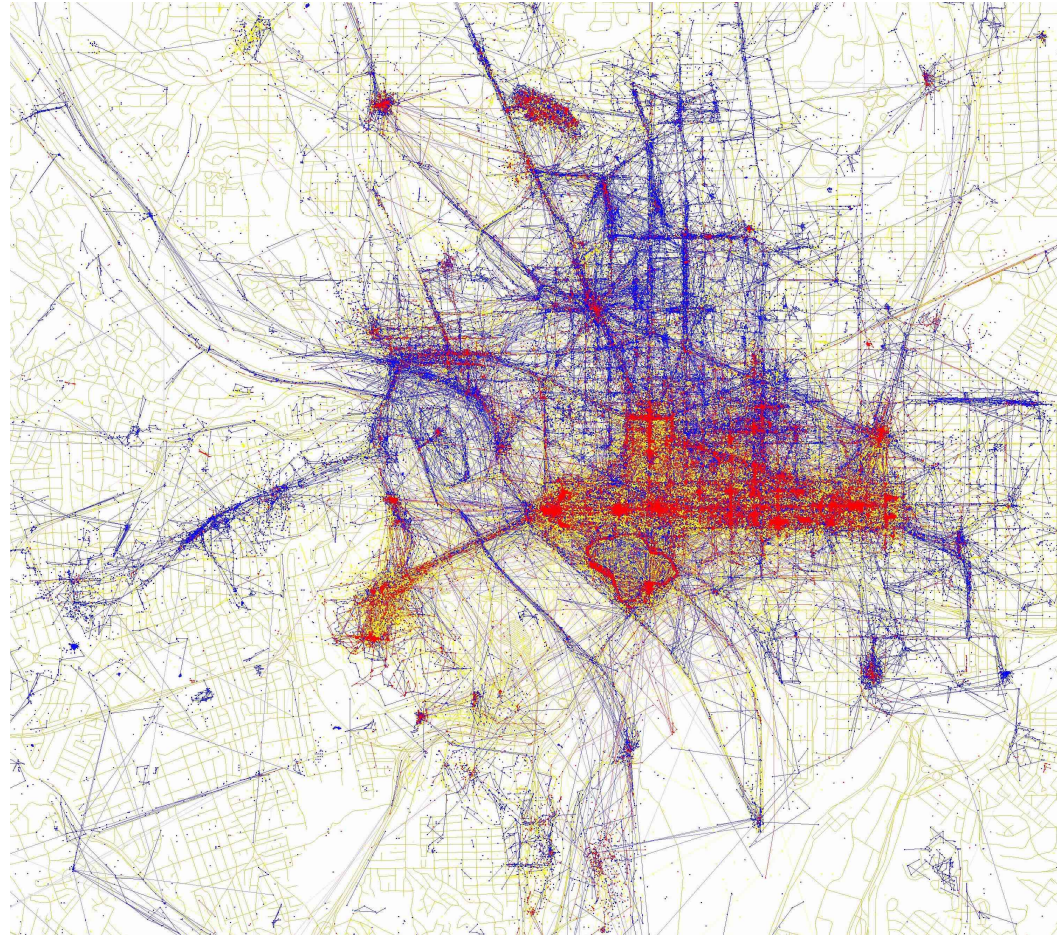
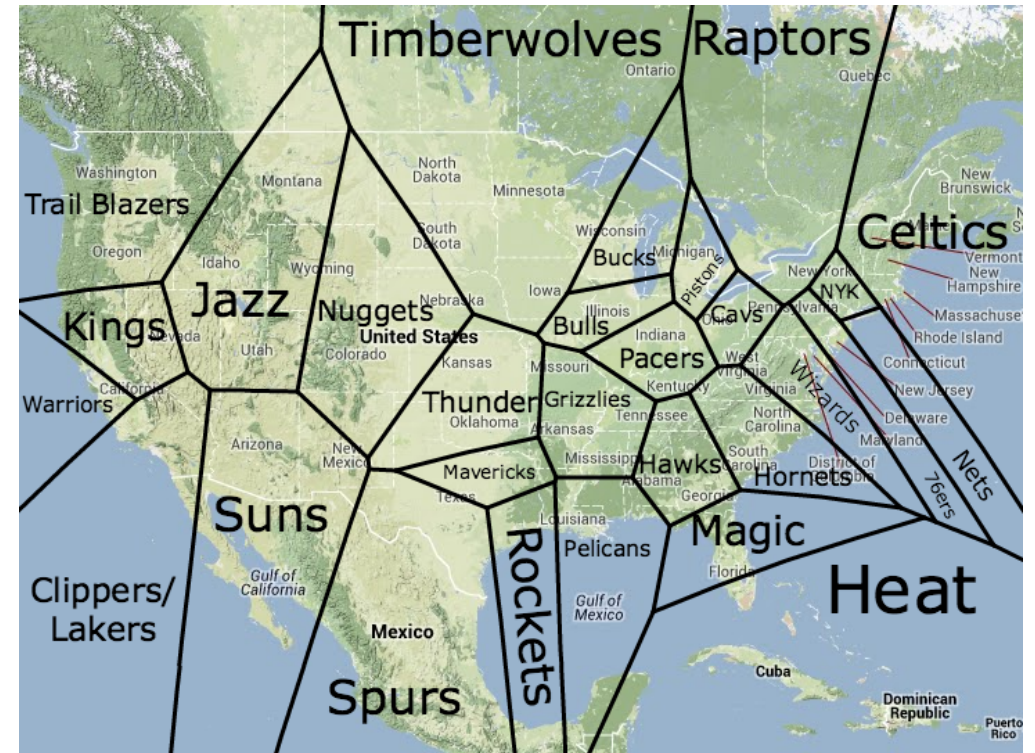

Locals vs. tourists in DC (blue = locals; red = tourists; yellow = unknown)

Voronoi maps

Voroni state boundaries, by Seth Kadish

Closest NBA teams

Maps with shapes

States Where Insured Could Renew Plans Before Change by Obama
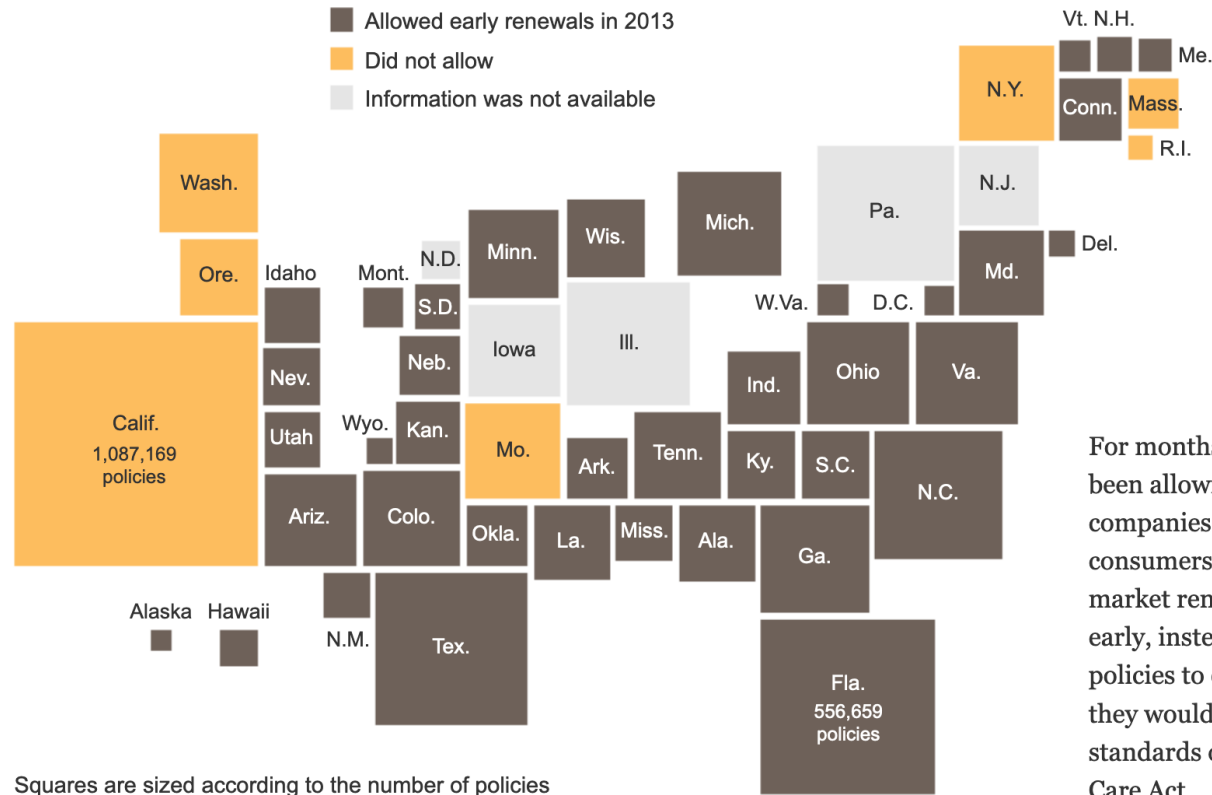
Squares are sized according to the number of policies in each state's individual insurance market in 2012.

For months, many states have been allowing insurance companies the option of letting consumers in the individual market renew their policies early, instead of waiting for the policies to expire in 2014, when they would not meet the standards of the Affordable Care Act.   Related Article »

Small multiples that look like maps

Seasonally Adjusted US Unemployment Rate 2000-2016

Data Source: bls.gov
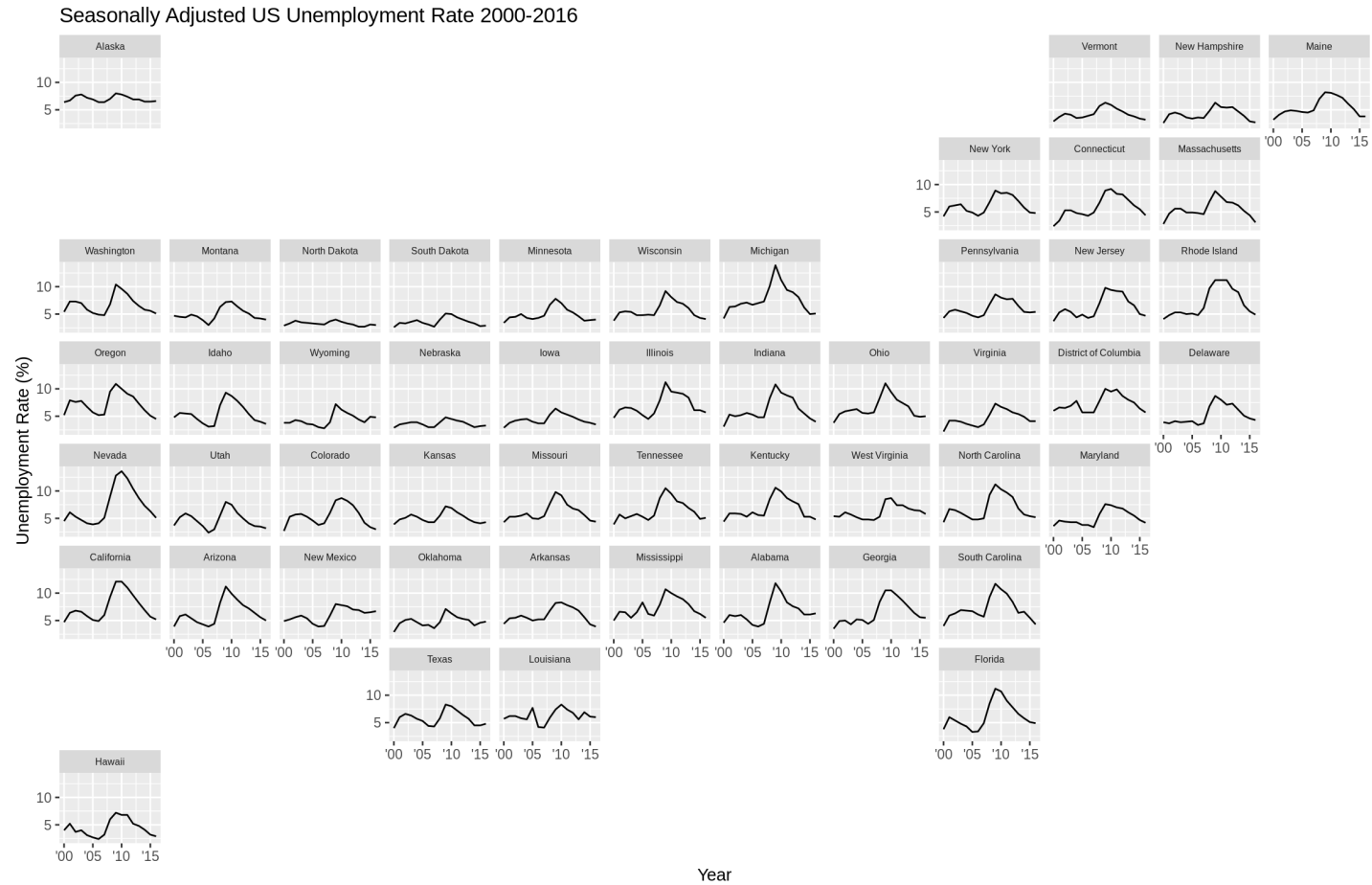
facet_geo() in the **geofacet** package

GIS in R with sf

Shapefiles

Geographic information is shared as shapefiles

These are *not* like regular single CSV files!

Shapefiles come as zipped files with
a bunch of different files inside

```
▼  📁 ne_110m_admin_0_countries
      📄 ne_110m_admin_0_countries.cpg
      📄 ne_110m_admin_0_countries.dbf
      📄 ne_110m_admin_0_countries.prj
      🦊 ne_110m_admin_0_...tries.README.html
      📄 ne_110m_admin_0_countries.shp
      📄 ne_110m_admin_0_countries.shx
      📄 ne_110m_admin_0_...untries.VERSION.txt
```
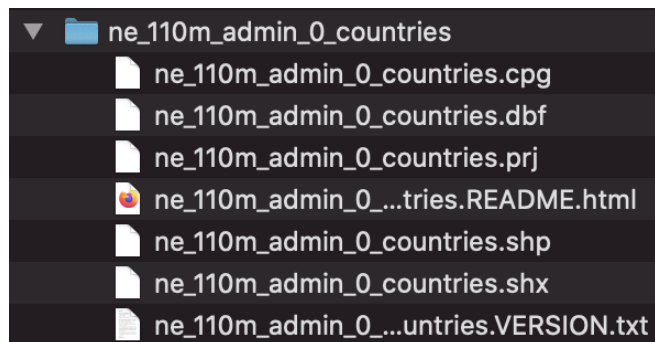
Structure of a shapefile

```
library(sf)
world_shapes <- read_sf("data/ne_110m_admin_0_countries/ne_110m_admin_0_countries.shp")
```

```
## Simple feature collection with 7 features and 3 fields
## geometry type:  MULTIPOLYGON
## dimension:      XY
## bbox:           xmin: -180 ymin: -18 xmax: 180 ymax: 83
## CRS:            4326
## # A tibble: 7 x 4
##   TYPE         GEOUNIT         ISO_A3                        geometry
##   <chr>        <chr>           <chr>               <MULTIPOLYGON [°]>
## 1 Sovereign … Fiji            FJI     (((180 -16, 180 -17, 179 -17, 179 -17, 179 -17, 179 …
## 2 Sovereign … Tanzania        TZA     (((34 -0.95, 34 -1.1, 38 -3.1, 38 -3.7, 39 -4.7, 39 …
## 3 Indetermin… Western Sahara  ESH     (((-8.7 28, -8.7 28, -8.7 27, -8.7 26, -12 26, -12 2…
## 4 Sovereign … Canada          CAN     (((-123 49, -123 49, -125 50, -126 50, -127 51, -128…
## 5 Country     United States … USA     (((-123 49, -120 49, -117 49, -116 49, -113 49, -110…
## 6 Sovereign … Kazakhstan      KAZ     (((87 49, 87 49, 86 48, 86 47, 85 47, 83 47, 82 46, …
## 7 Sovereign … Uzbekistan      UZB     (((56 41, 56 45, 59 46, 59 46, 60 45, 61 44, 62 44, …
```

Where to find shapefiles

Natural Earth for international maps

US Census Bureau for US maps

For anything else...

Google

🔍 shapefiles for _____ ✕

Scales

1:10m = 1:10,000,000
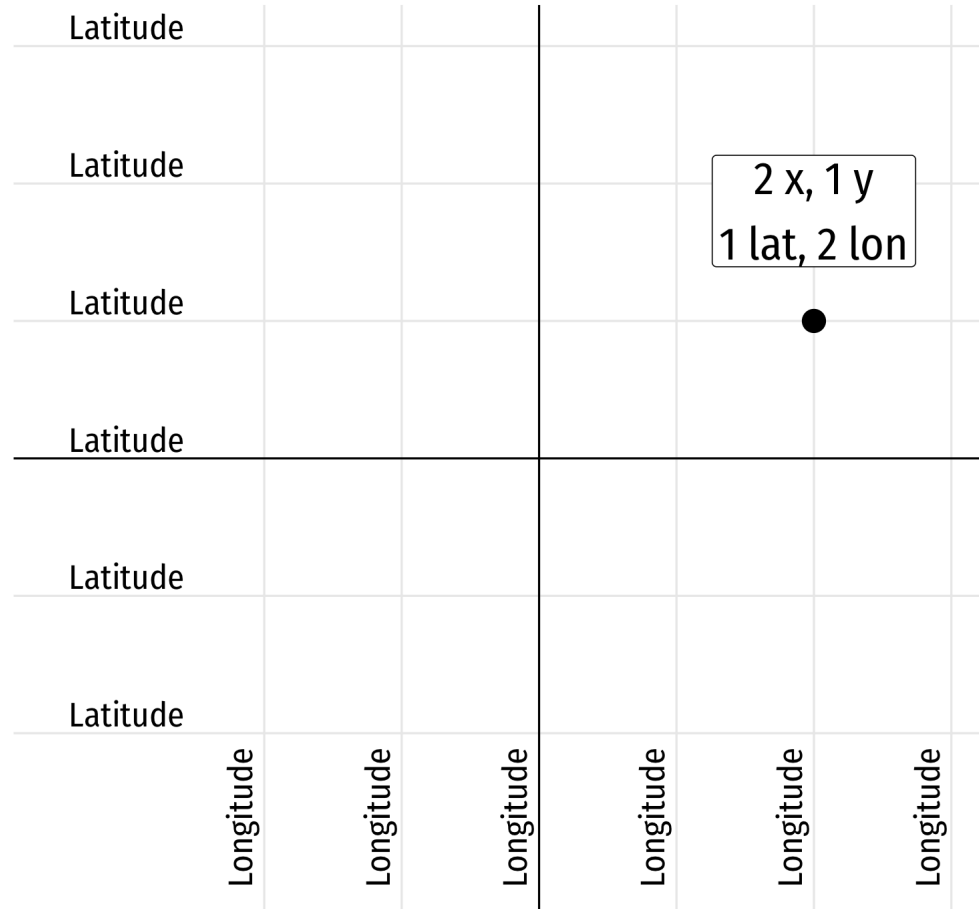
1 cm = 100 km

1:50m = 1:50,000,000

1cm = 500 km

1:110m = 1:110,000,000

1 cm = 1,100 km

Using too high of a resolution makes your maps slow and huge

Latitude and longitude
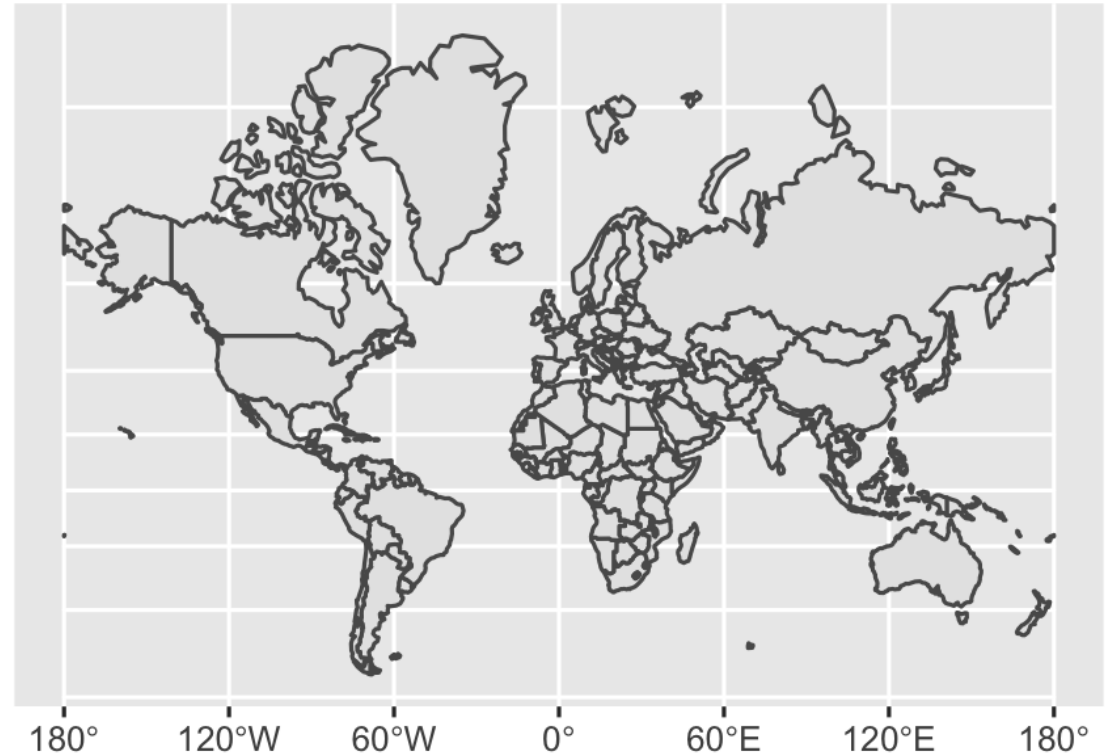
The magic `geometry` column

As long as you have a magic `geometry` column, all you need to do to plot maps is `geom_sf()`

```
ggplot() +
  geom_sf(data = world_shapes)
```

The magic `geometry` column

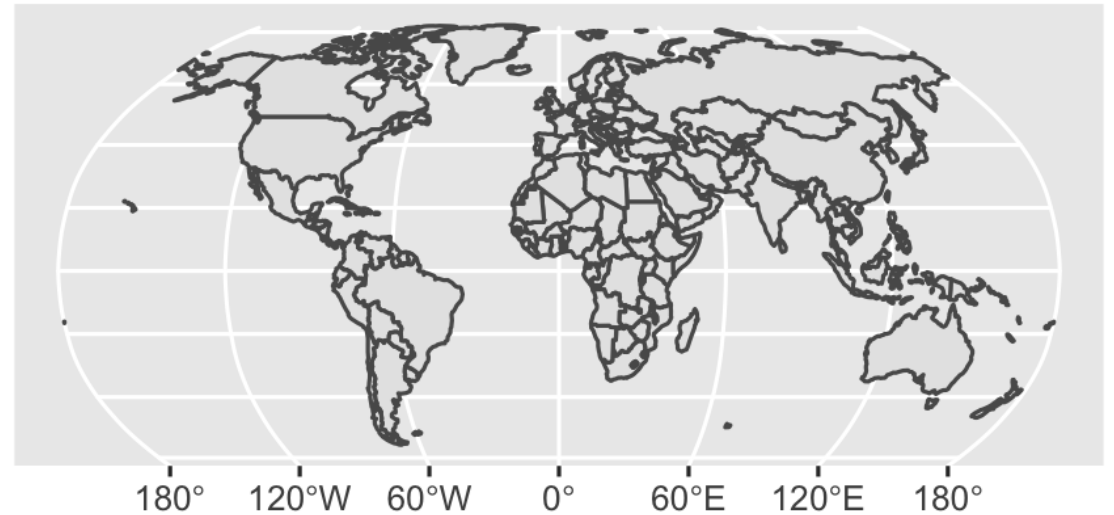Use `coord_sf()` to change projections

```
ggplot() +
  geom_sf(data = world_shapes) +
  coord_sf(crs = "+proj=merc")
```

The magic `geometry` column

Use `coord_sf()` to change projections

```
ggplot() +
  geom_sf(data = world_shapes) +
  coord_sf(crs = "+proj=robin")
```

Use aesthetics like normal

All regular ggplot layers and aesthetics work

```
ggplot() +
  geom_sf(data = world_shapes,
          aes(fill = POP_EST),
          color = "white", size = 0.15) +
  coord_sf(crs = "+proj=robin") +
  scale_fill_gradient(labels = scales::comma)
  labs(fill = NULL) +
  theme_void() +
  theme(legend.position = "bottom")
```


500,000,000,000,000
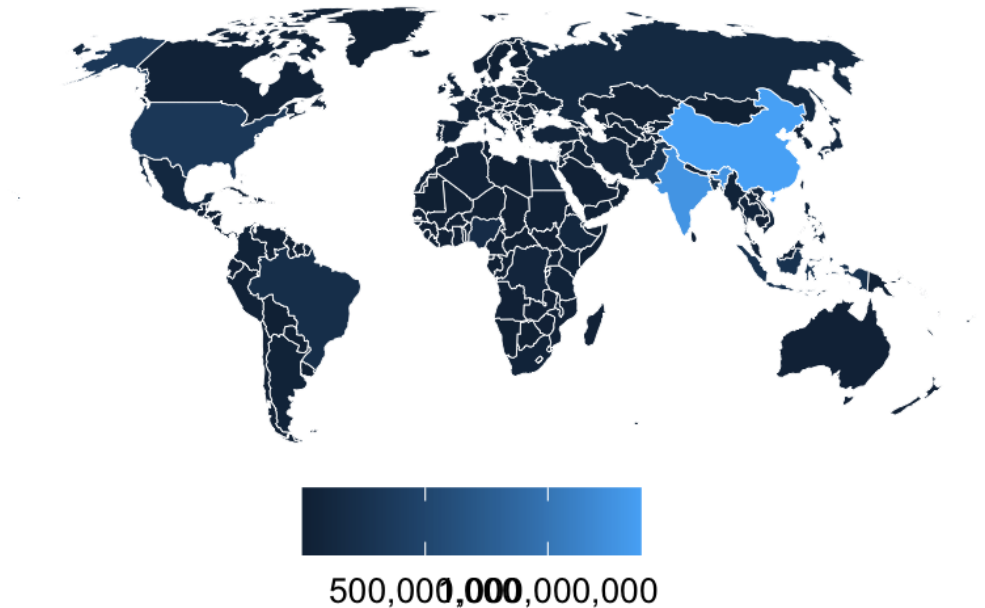
No geometry column?

Make your own with `st_as_sf()`

```
other_data
```

```
## # A tibble: 2 x 3
##   city            long   lat
##   <chr>          <dbl> <dbl>
## 1 Atlanta        -84.4  33.8
## 2 Washington, DC -77.1  38.9
```

```
other_data %>%
  st_as_sf(coords = c("long", "lat"),
           crs = 4326)
```

```
## Simple feature collection with 2 features and 1 field
## geometry type:  POINT
## dimension:      XY
## bbox:           xmin: -84 ymin: 34 xmax: -77 ymax: 39
## CRS:            EPSG:4326
## # A tibble: 2 x 2
##   city               geometry
##   <chr>           <POINT [°]>
## 1 Atlanta           (-84 34)
## 2 Washington, DC    (-77 39)
```

sf is for all GIS stuff

Draw maps

Calculate distances between points

Count observations in a given area

Anything else related to geography!

See here or here for full textbooks

`geom_sf()` is today's standard

You'll sometimes find older tutorials and StackOverflow answers about using `geom_map()` or **ggmap** or other things

Those still work, but they don't use the same magical **sf** system with easy-to-convert projections and other GIS stuff

Stick with **sf** and `geom_sf()` and your life will be easy