



NYU | TANDON

# Information/Data Visualization

Lecture 3 –  
2D Visualization – Spatial Data

Qi Sun  
[qisun@nyu.edu](mailto:qisun@nyu.edu)

# Logistics

- Clarification on final project options
  - [1] Your research project with visualization component
  - [2] Reproducing an IEEE VIS/TVCG (or others with discussion) paper
    - <https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=2945>
  - [3] Real-world applications (see examples in Meta Info on Brightspace)
- Deliverables: Presentation + Code + Report

# Logistics

- First guest lecture (3/6)
  - Gromit Chan, Adobe Research, San Jose, CA
  - Visualization and data management, including large scale data visualization



# Geographical Data

**Any data that contains geographical information**

## Spatial objects

- Counties
- Regions
- Buildings
- Rivers, lakes
- Etc.

## Non-spatial geolocated objects

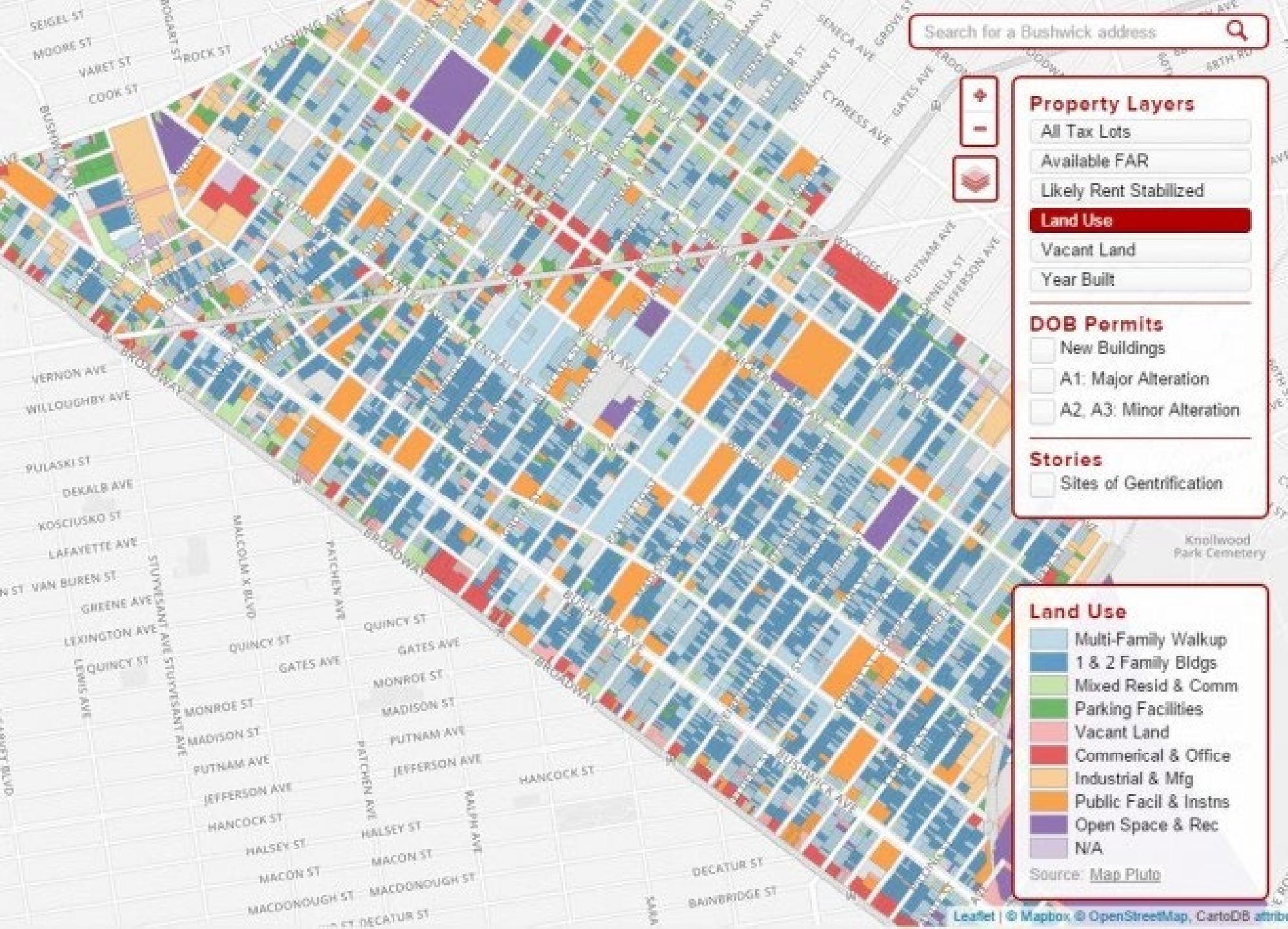
- Cars
- People
- Animals
- Weather stations
- Etc.

# Spatial Objects

- Shape, spatial extent matters.
- Map is the main object of interest.



Roads to Rome: <https://benedikt-gross.de/projects/roads-to-rome>

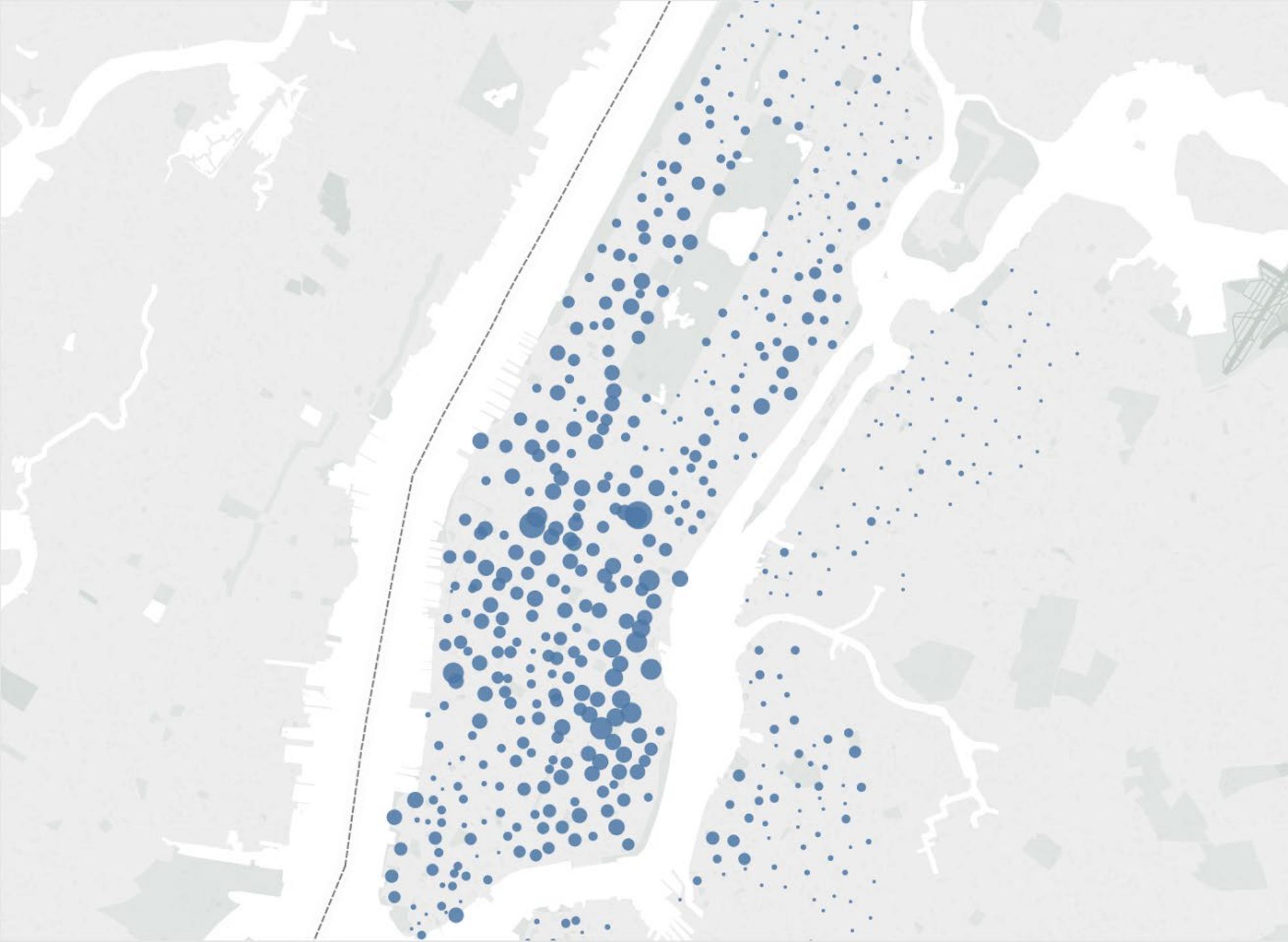


Source: Am I Rent Stabilized?  
Data Visualization as Housing Activism with Chris Hendrick

<http://urbandemos.nyu.edu/am-i-rent-stabilized-data-visualization-as-housing-activism-with-chris-hendrick/>

# Non-Spatial Objects

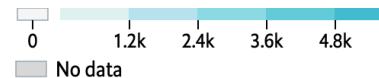
- Shape does not matter.
- Map is used as a reference.



Number of citi bikes in NYC.  
Credit: Erico Bertini

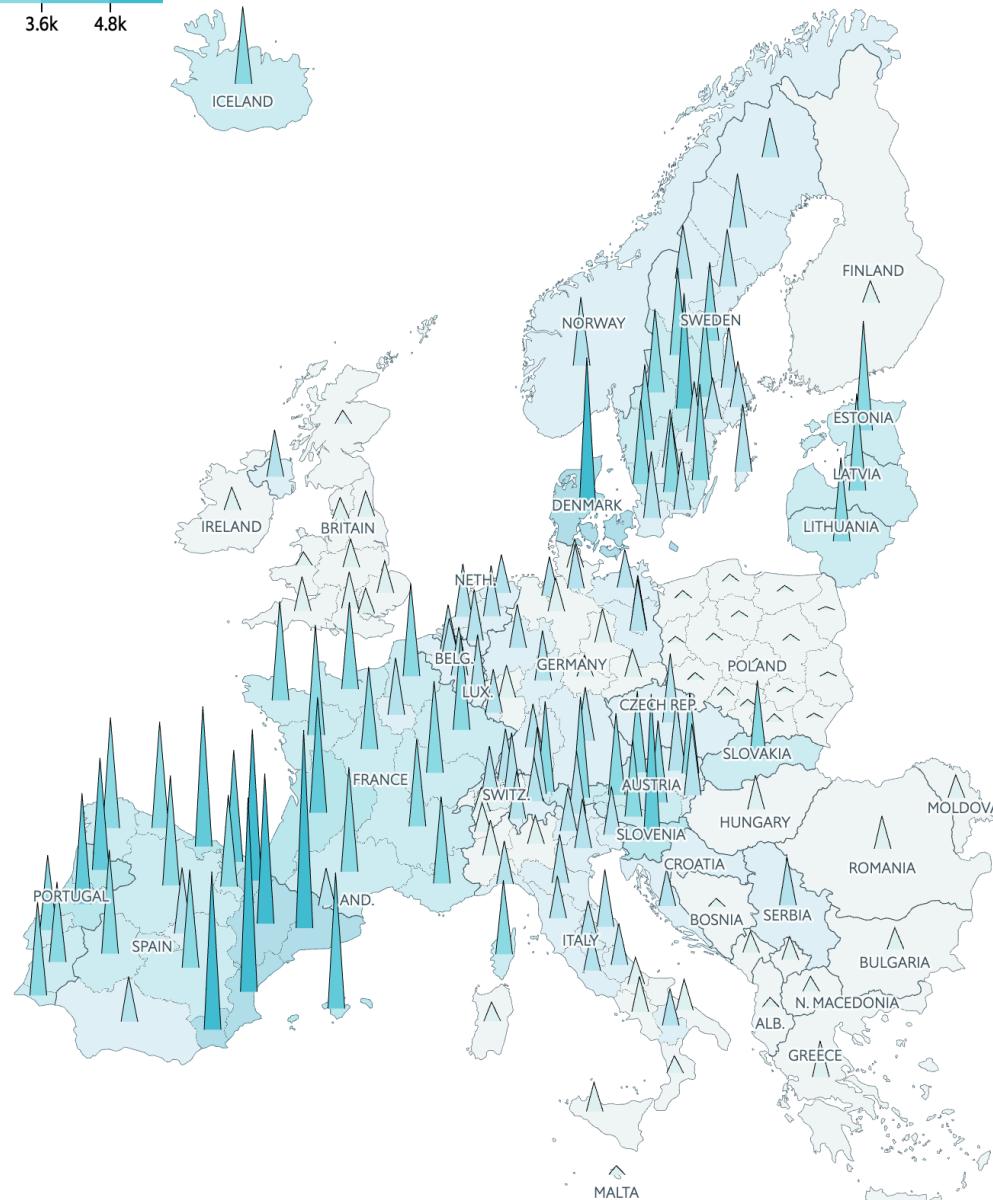
**Covid-19 in Europe**

Per 100,000 people



Cases, last 7 days

Deaths, last 7 days

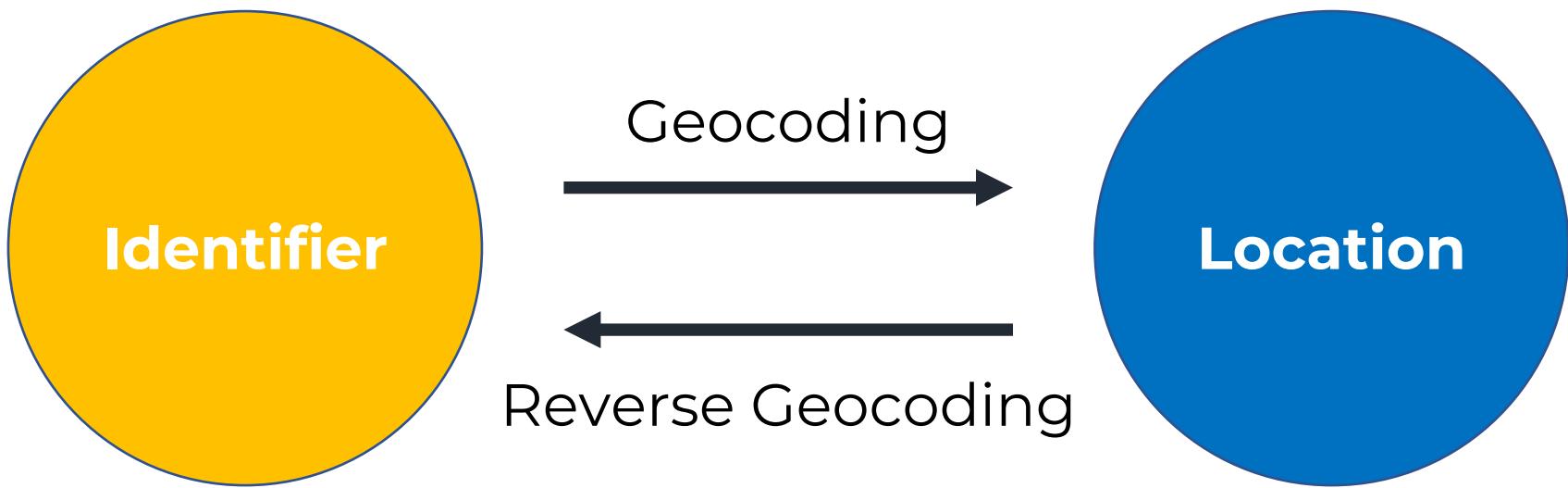


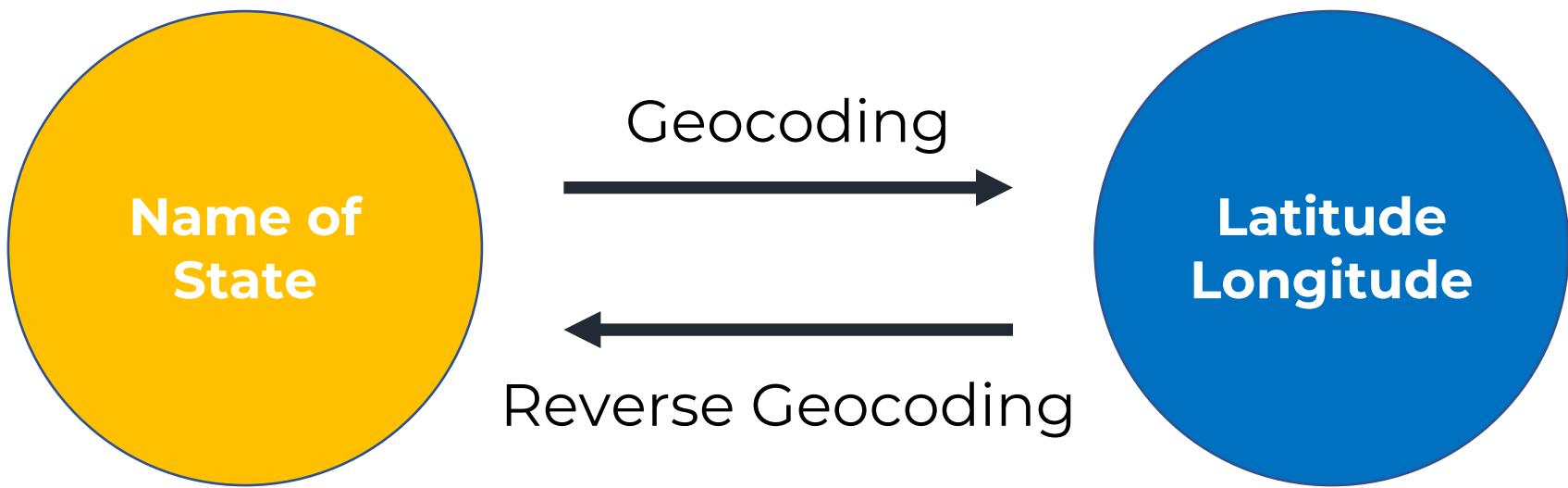
**Tracking the coronavirus across Europe**  
The Economist, Feb 5, 2022

<https://www.economist.com/graphic-detail/tracking-coronavirus-across-europe>

# Types of geographical data

- Regions (e.g. country borders)
- Locations (e.g. latitude, longitude)
- Identifiers (e.g. zip code, street name, borough)



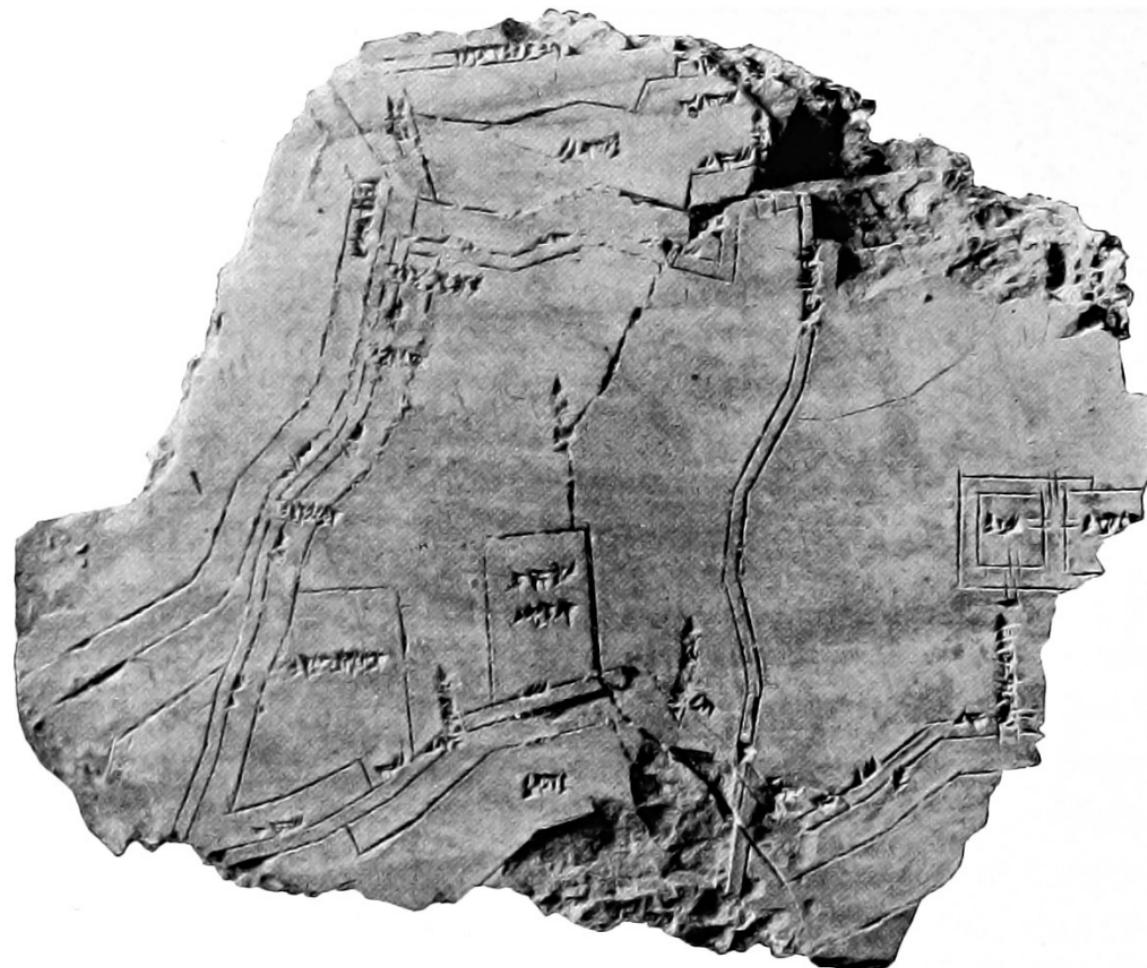


What is a common way to visualize  
geographical data?

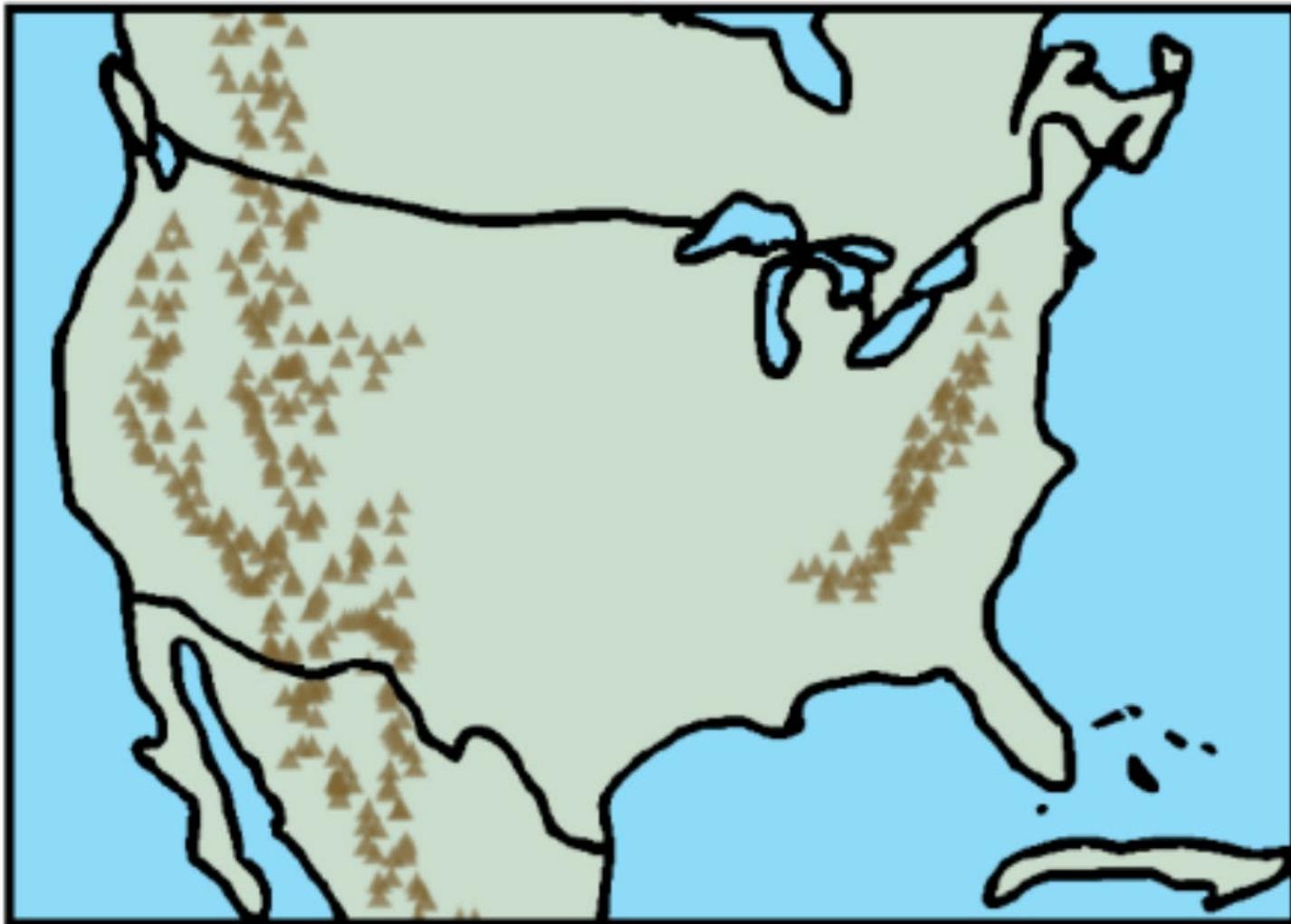
Map!

# What is a map?

**Clay tablet with map of the  
Babylonian city of Nippur  
(ca. 1400 BC)**



# What is a map?



# What is a map?

Dimensionality reduction

3D



2D

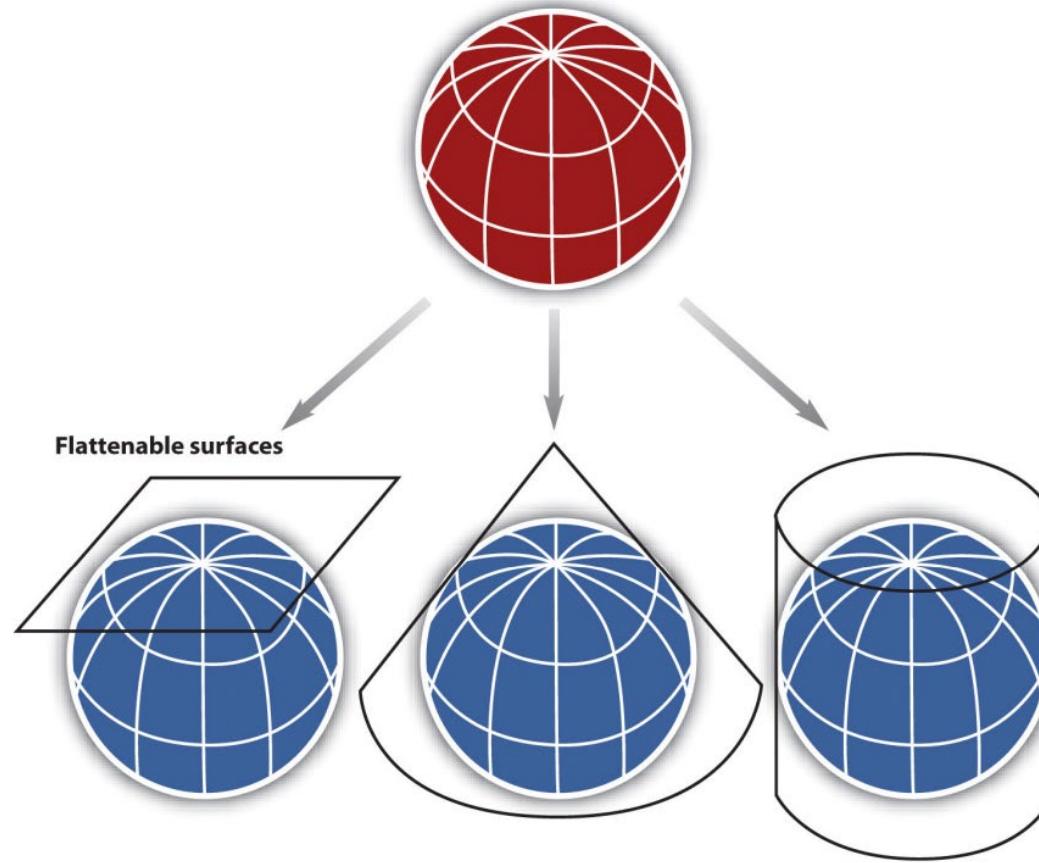


credit: Prof. Michelle Borkin

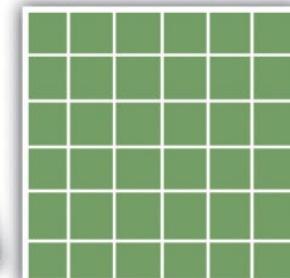
# Map Projections

# What is a map?

## Dimensionality Reduction



Flat maps

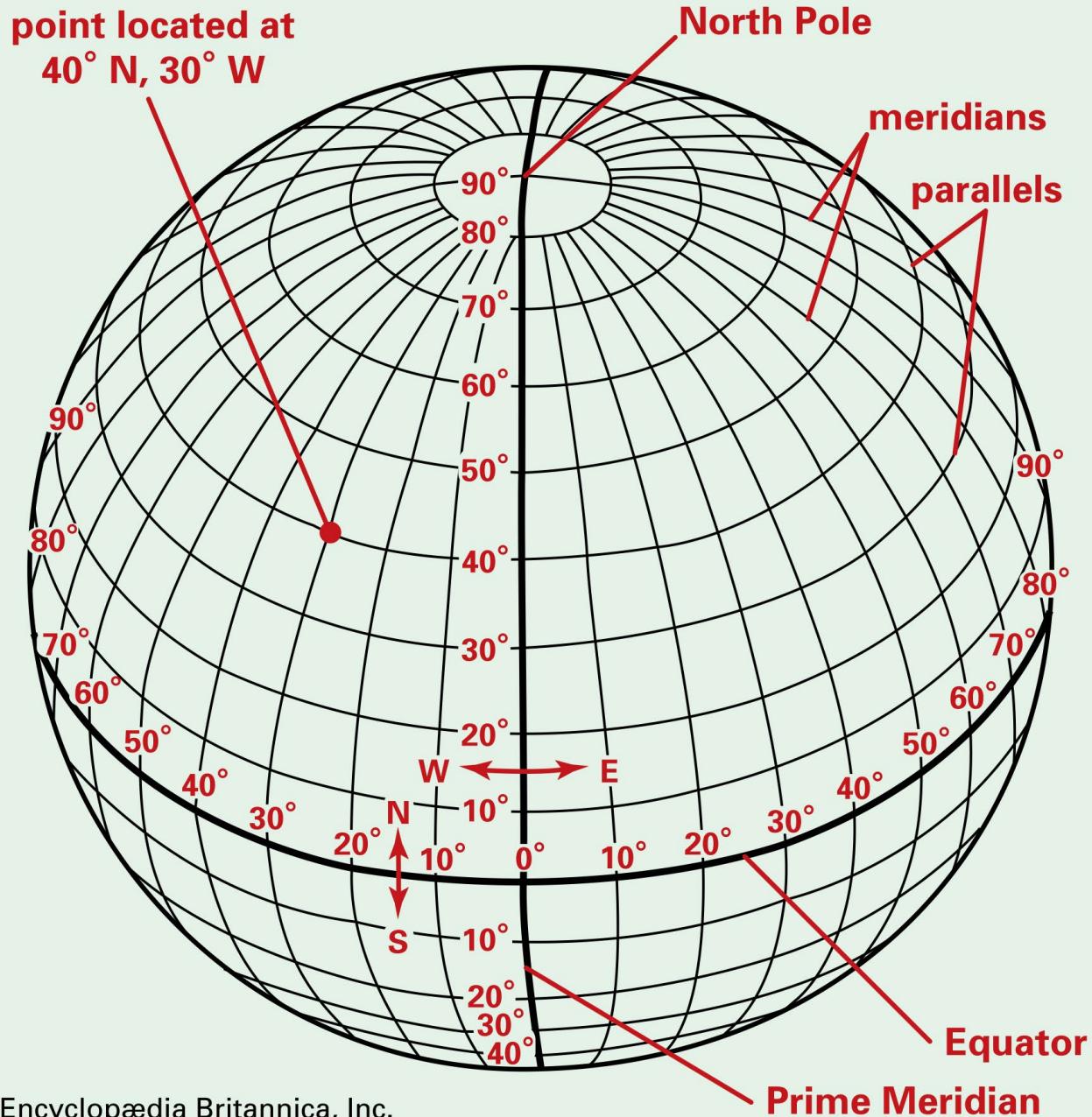


azimuthal

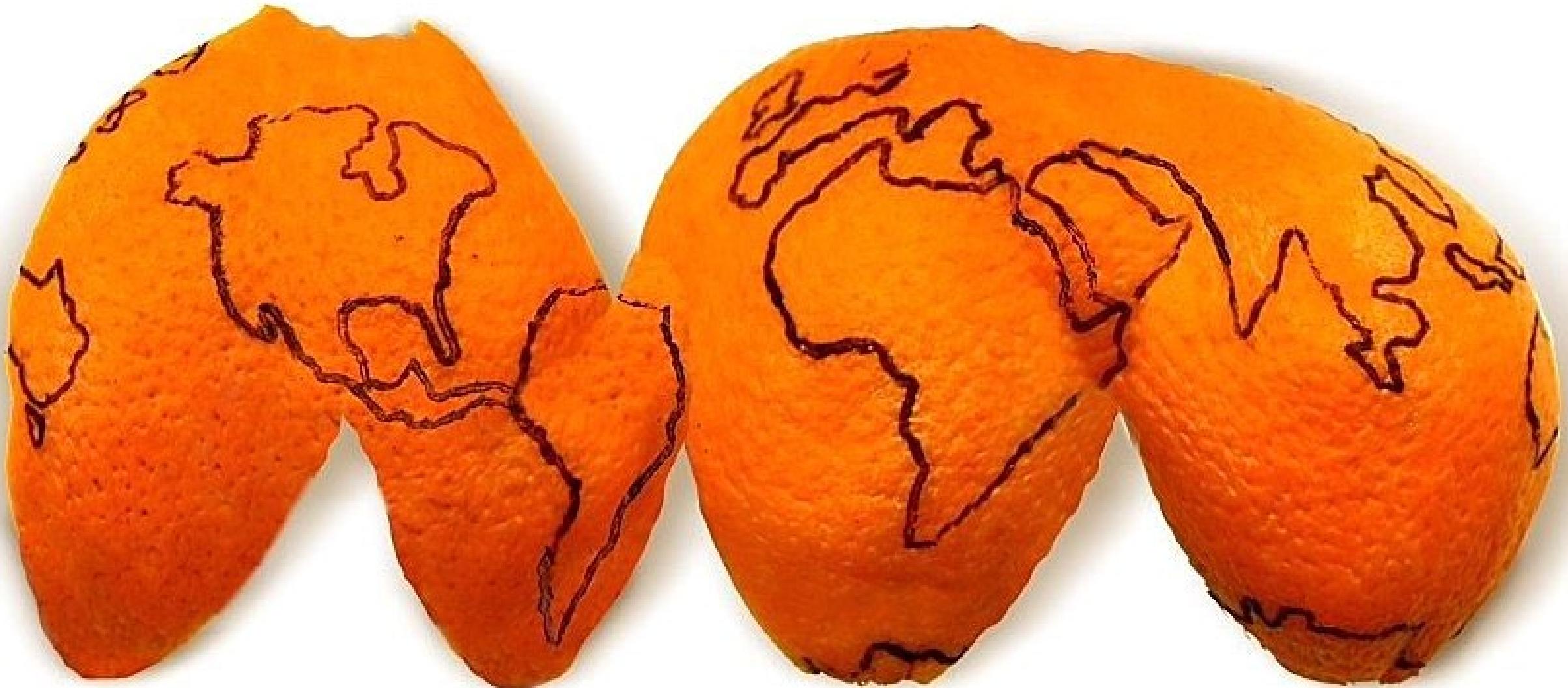
conic

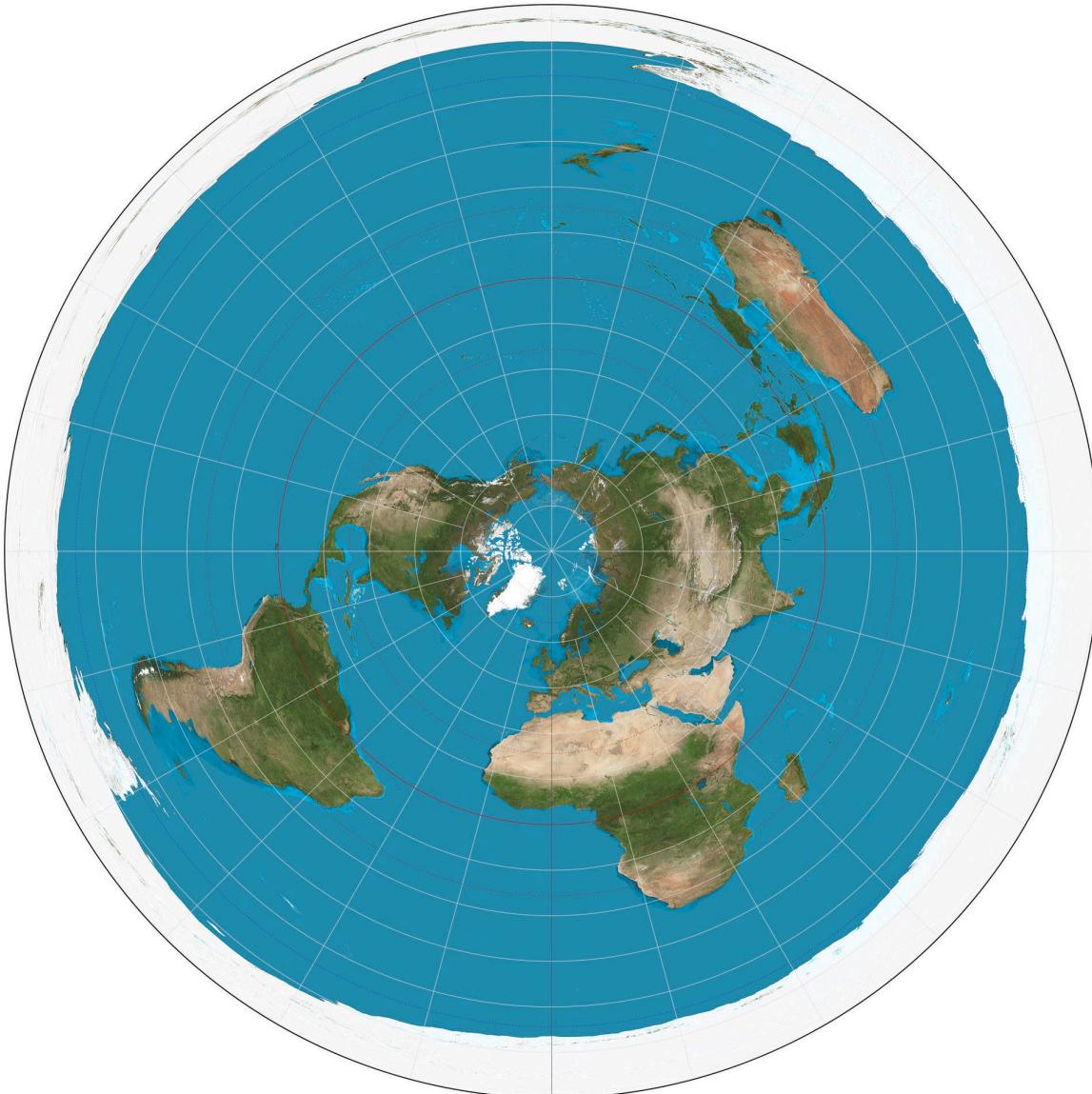
cylindrical

## PARALLELS OF LATITUDE AND MERIDIANS OF LONGITUDE



All projections involve **cut** and **distortion**!

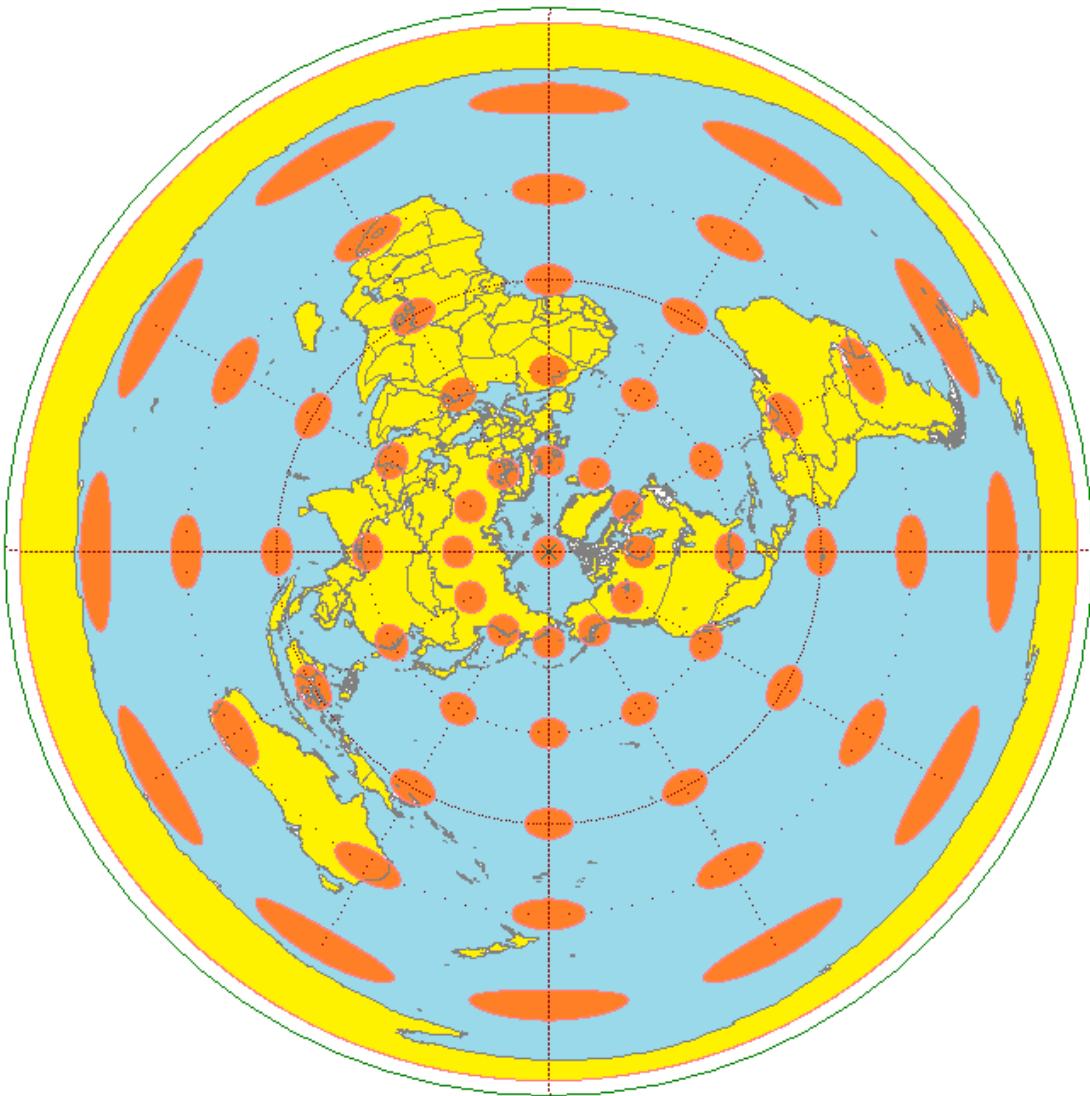




## Azimuthal Equidistant

Preserves distance and direction from center point.

Used for travel, propagation from center point.

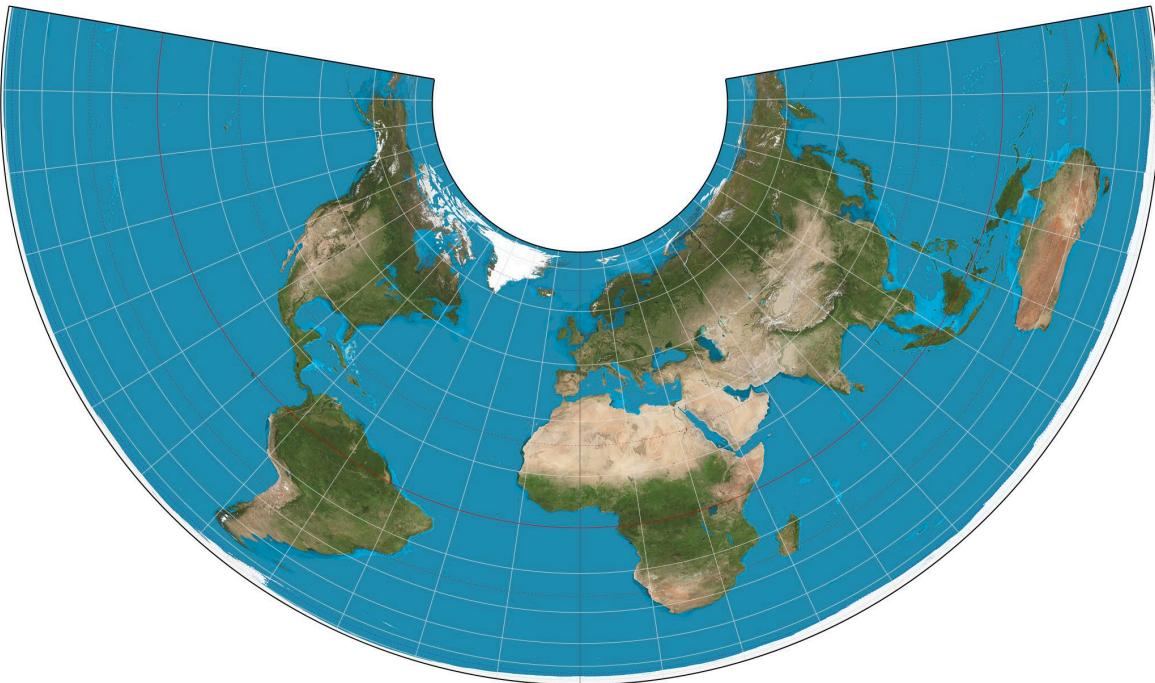


Tissot's indicatrix

## Azimuthal Equidistant

Preserves distance and direction from center point.

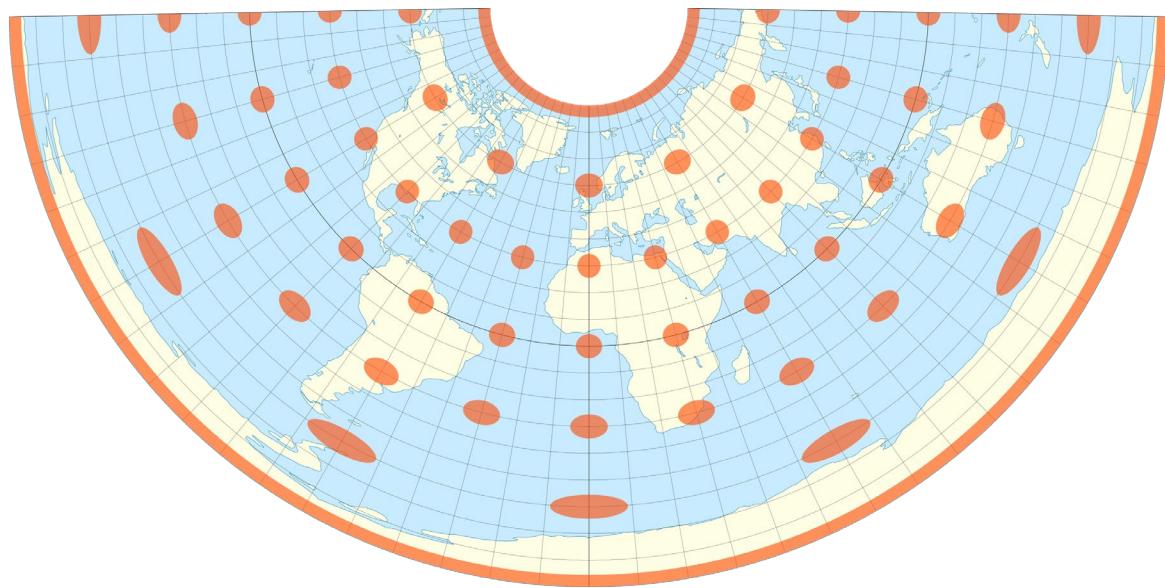
Used for travel, propagation from center point.



## **Equidistant Conic**

Preserves proportional area of regions.

Used for land survey, choropleth map.

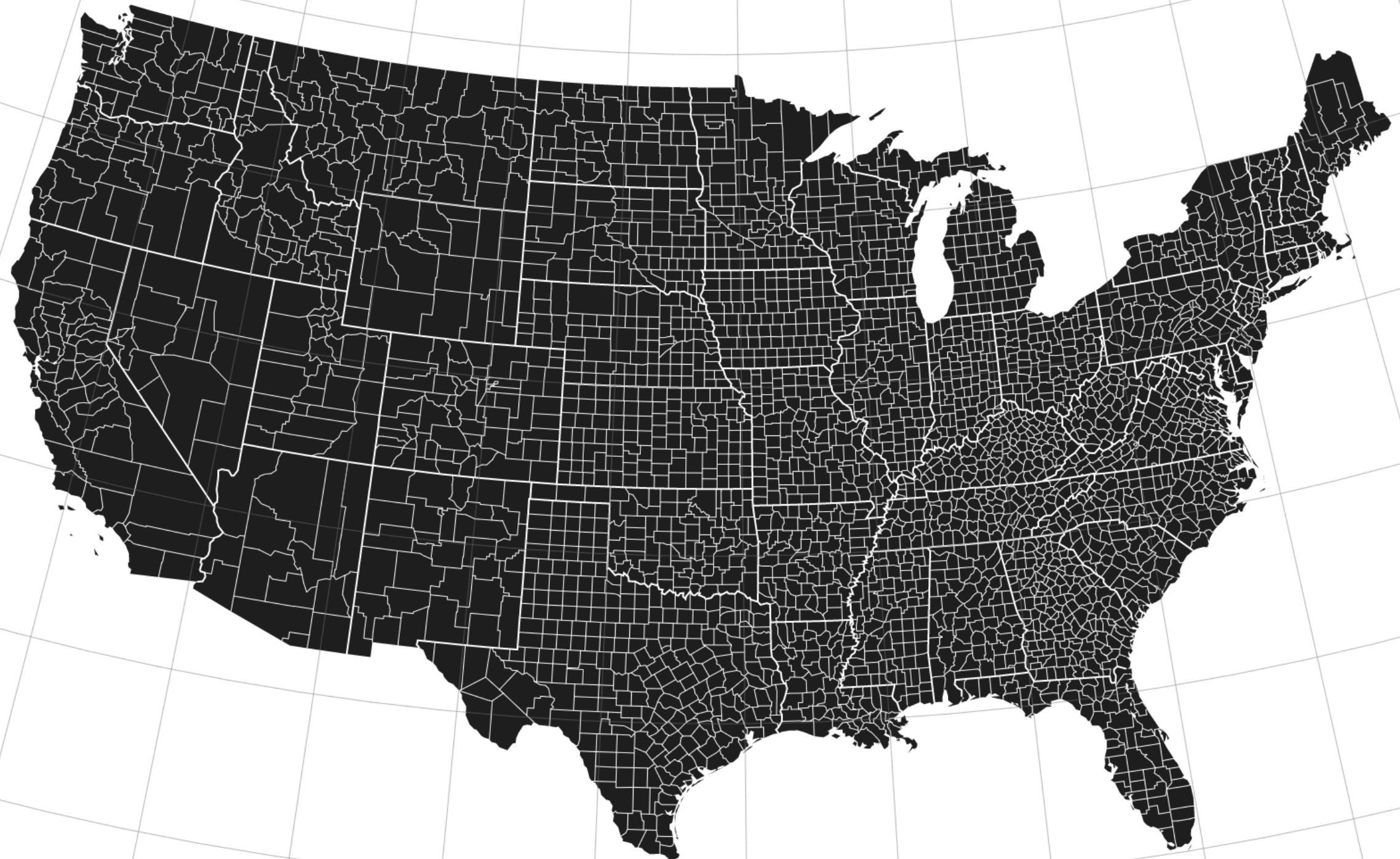


Tissot's indicatrix

## **Equidistant Conic**

Preserves proportional area of regions.

Used for land survey, choropleth map.





## Mercator

Preserves angle (conformal).

Used for navigation.

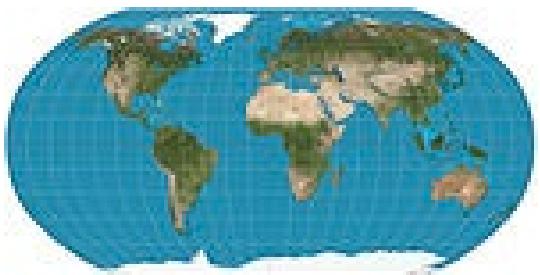


Tissot's indicatrix

## Cylindrical Mercator

Preserves angle (conformal).

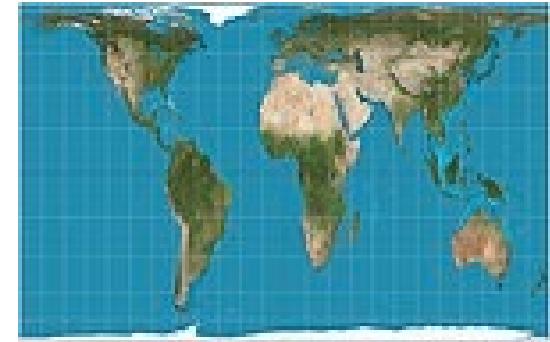
Used for navigation.



Robinson



Goode homolosine



Gall-Peters



Orthographic



Dymaxion map



Waterman butterfly

# Projection

- Quantities to preserve (distance, area, angle, etc.)
- Different ways of cutting open a sphere.
- Many projection maps are available.
  - The most popular one is web Mercator from Google.

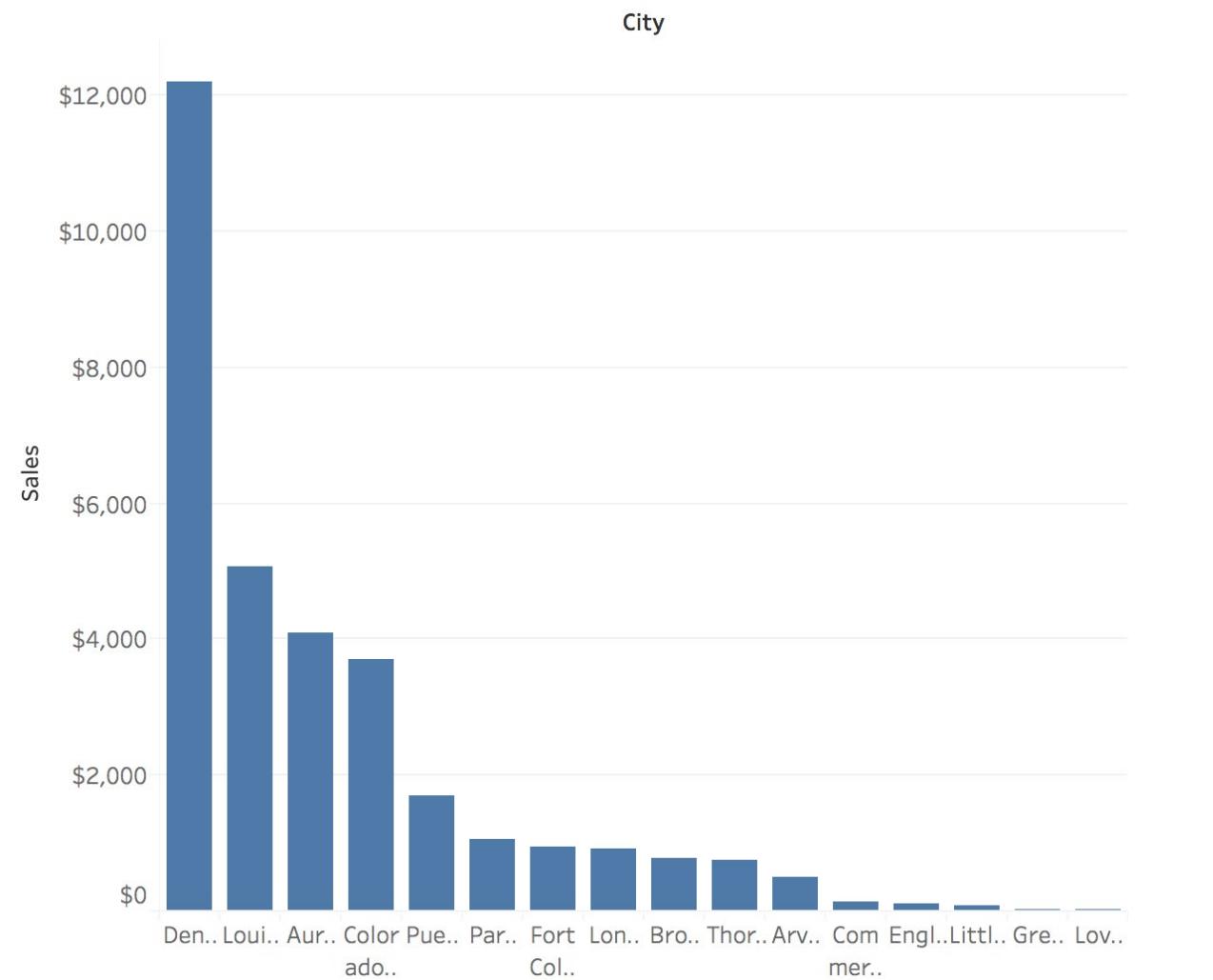
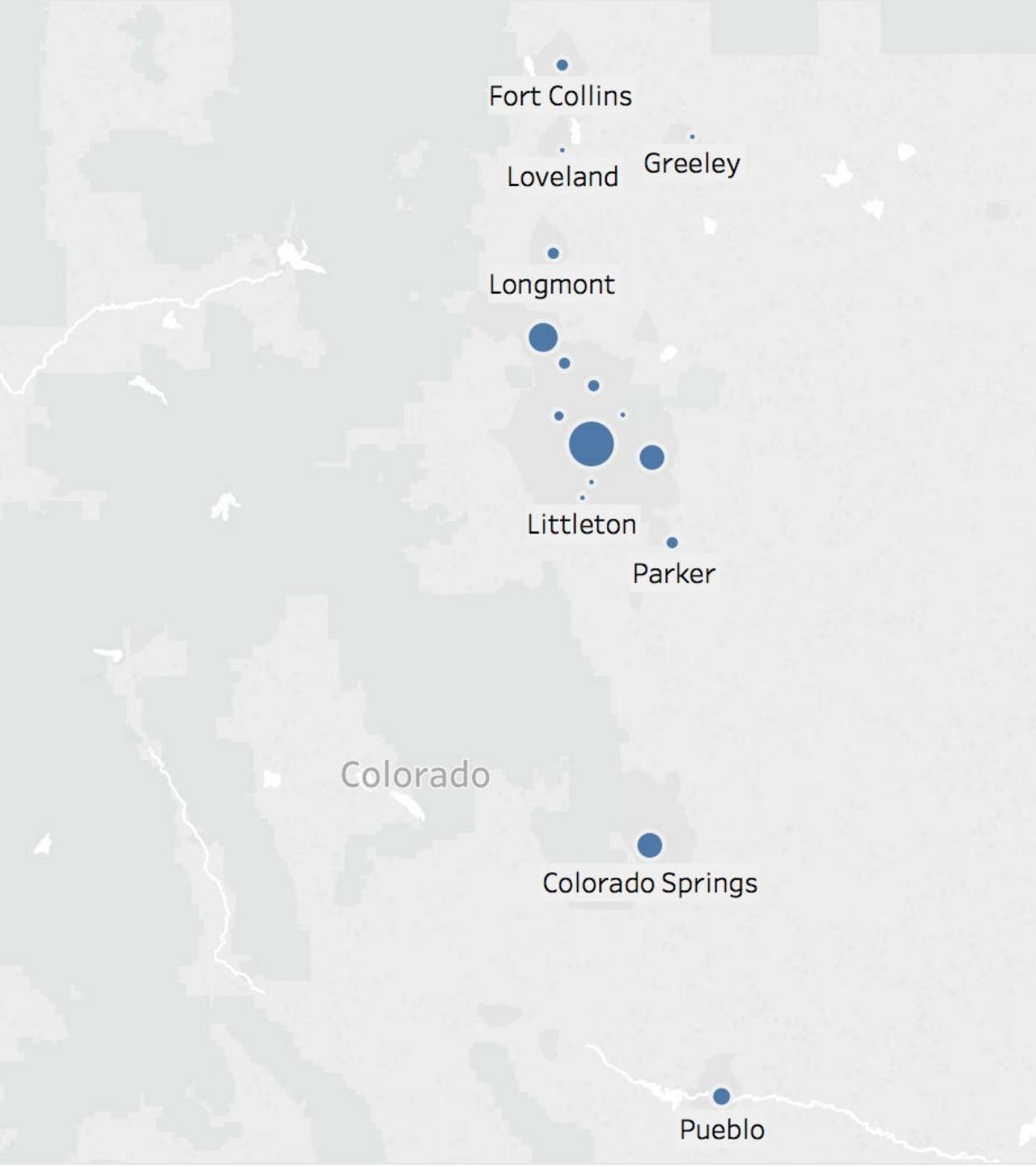
# When/why use a map?

Fundamental question every time you have to  
choose how to visualize geographical data.

# Limitation of Map

Maps use *space* to represent *space*.

Space *cannot* be used to encode other types of information.



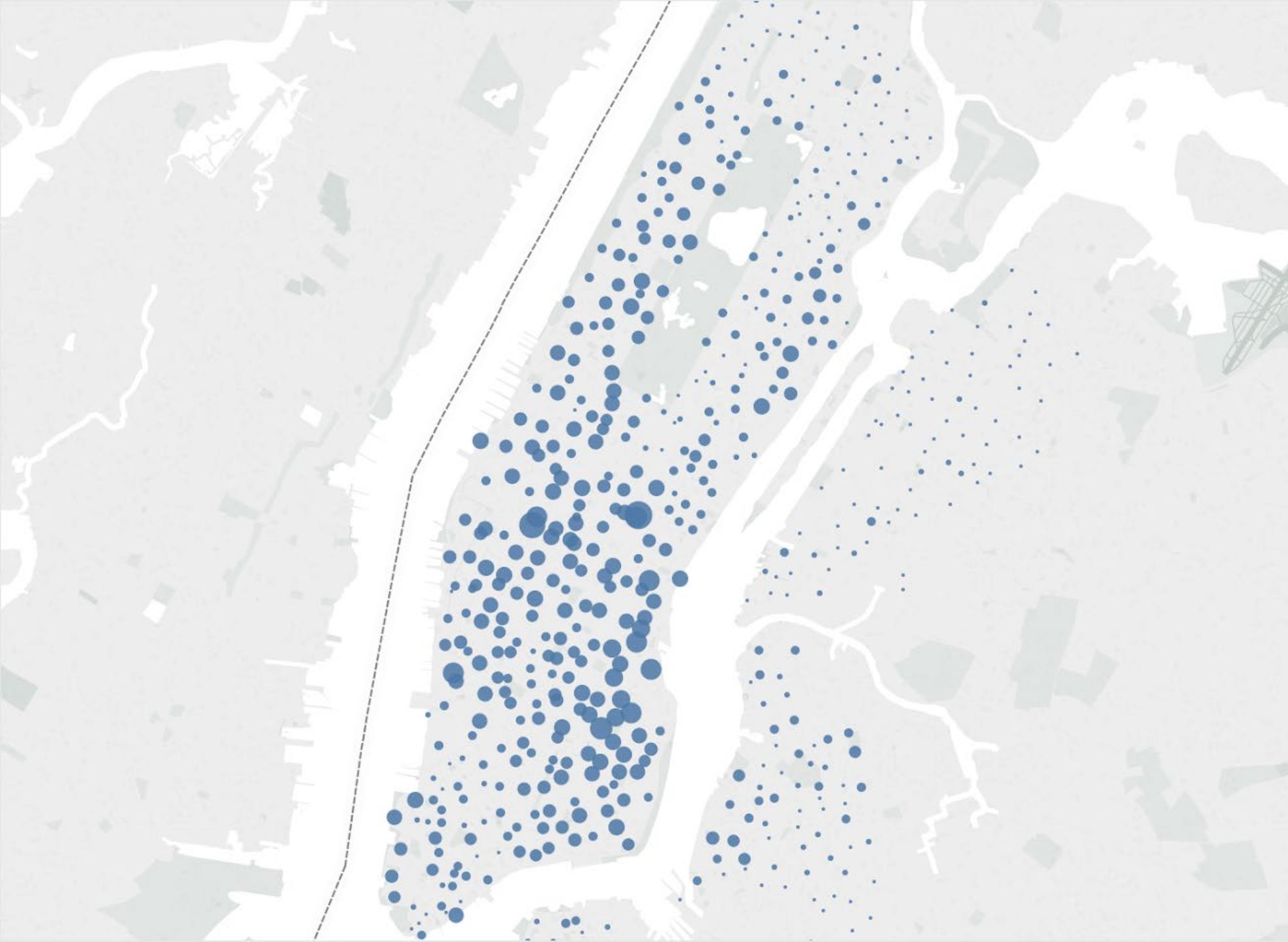
Sales in cities of Colorado  
Credit: Erico Bertini

# When to use maps?

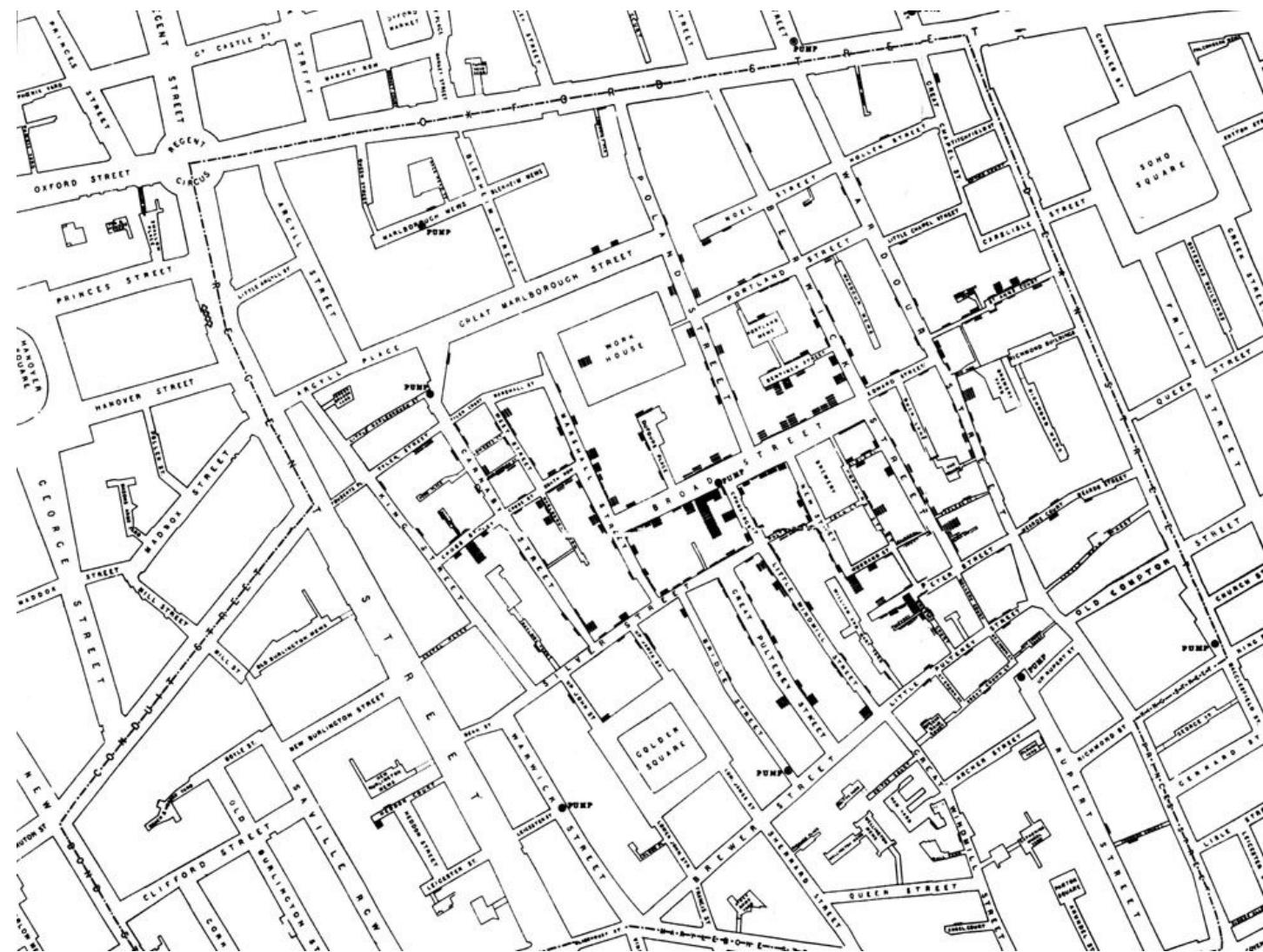
1. When the question is inherently spatial.
2. When map helps find information needed.

# When to use maps?

1. When the question is inherently spatial.
  - a) Question that correlates a phenomenon to spatial location/objects.
  - b) Questions pertaining phenomenon for which spatial proximity or extend is relevant.



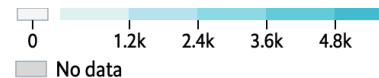
Number of citi bikes in NYC.  
Credit: Erico Bertini



London cholera map  
Credit: John Snow, 1854

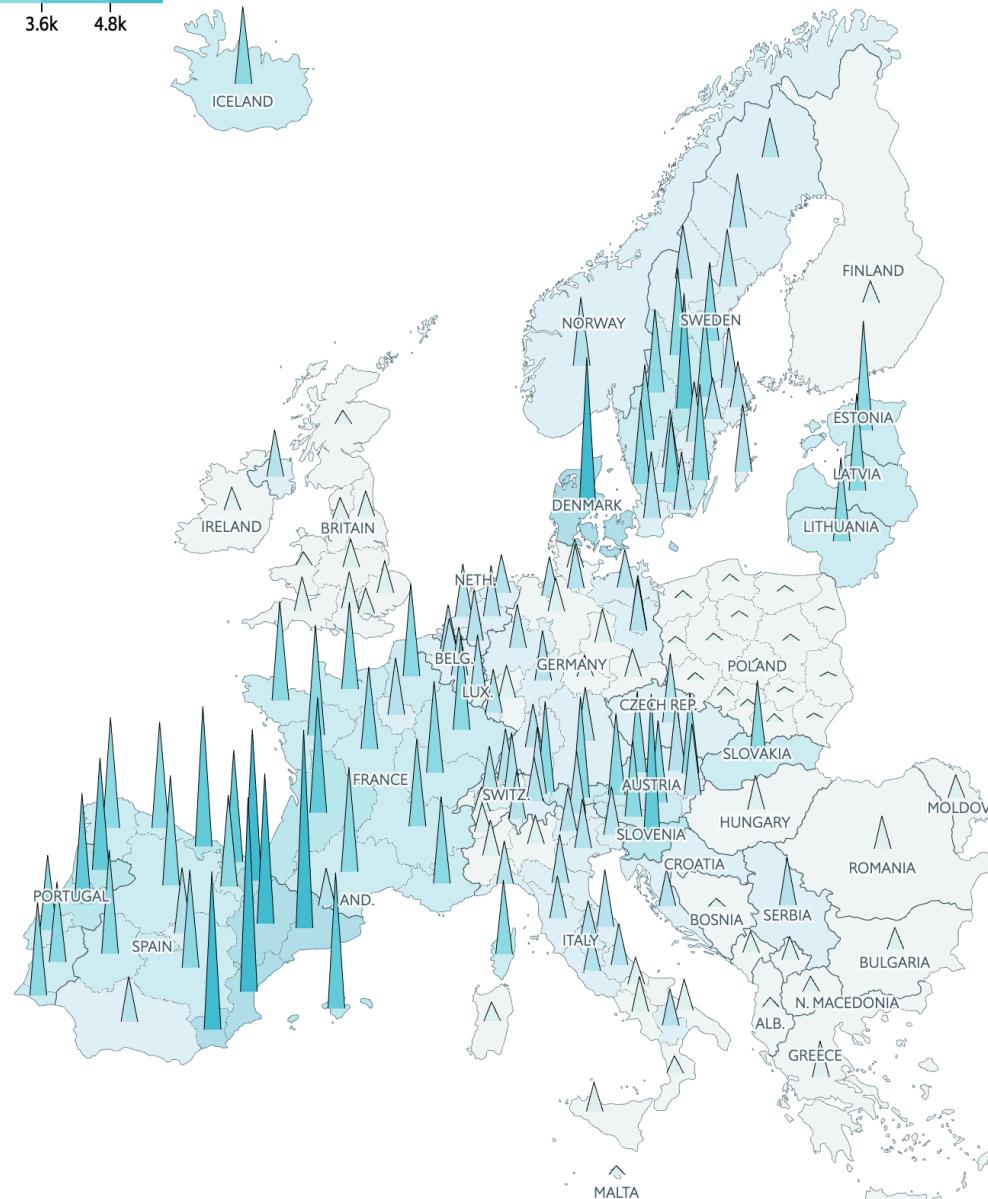
**Covid-19 in Europe**

Per 100,000 people



Cases, last 7 days

Deaths, last 7 days



**Tracking the coronavirus across Europe**  
The Economist, Feb 5, 2022

<https://www.economist.com/graphic-detail/tracking-coronavirus-across-europe>

## Latest on Data visualisation



Data visualisation: how the FT newsroom designs maps



How to lie with maps



Dump the PowerPoints or risk losing money



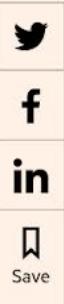
How to navigate data's pink and blue problem

The Chart Doctor **Data visualisation**

+ Add to myFT

## A love of maps should mean using fewer to illustrate data better

When one map is not enough, and when you should use none at all



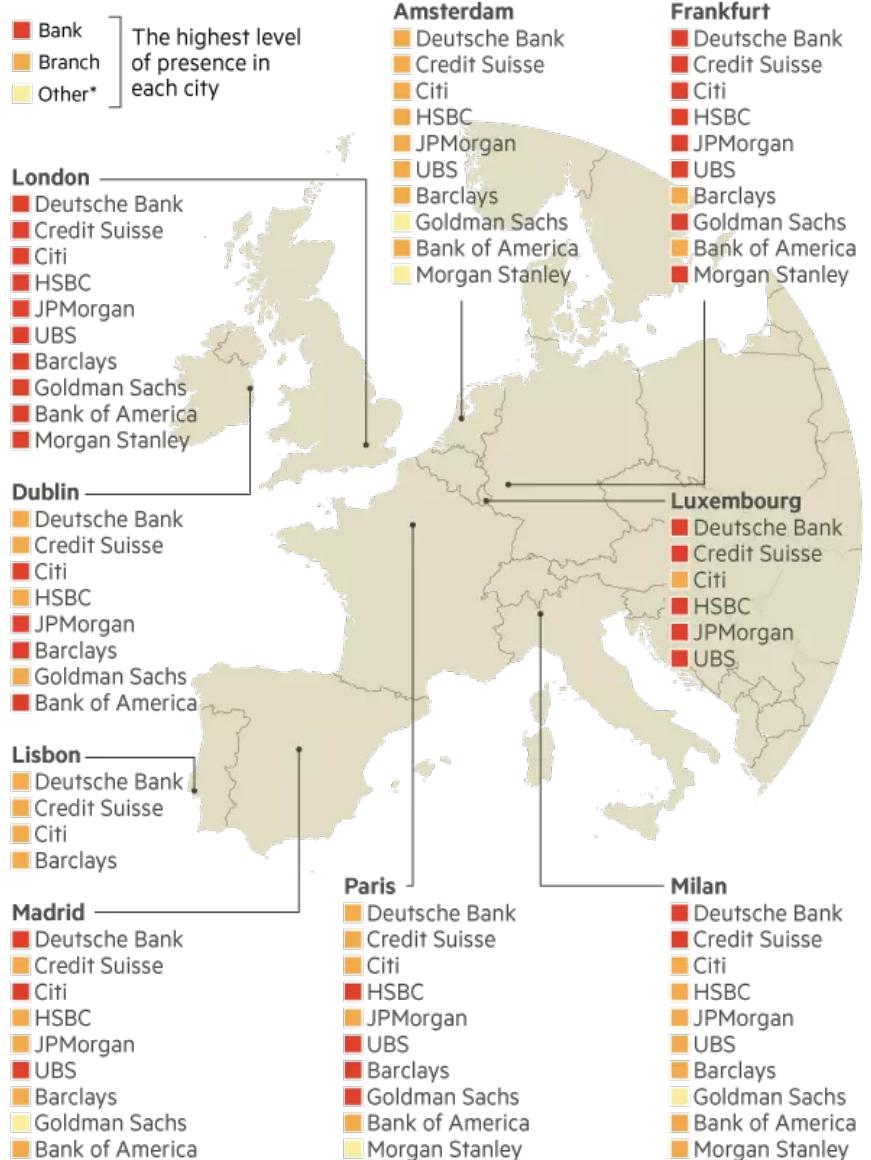
THE  
**CHART**  
DOCTOR

## When maps are the wrong route

However, maps are not the solution to every visualisation problem with geographical data. As the [FT Visual Vocabulary](#) highlights, “spatial” is just one of many possible relationships in data and maps and should be used “only when specific locations or geographical patterns in data are more important to the reader than anything else”. Ignore this advice and maps may end up leaving readers lost.

Recently, some painstaking research by Laura Noonan, the FT’s investment banking correspondent, produced a data set of top banks’ presence in eight cities lining up to take London’s role as the EU’s banking centre. Nearly 5,000 entities, each allocated to one of the eight cities.

## The Brexit banking matrix: The contenders lining up for London's crown

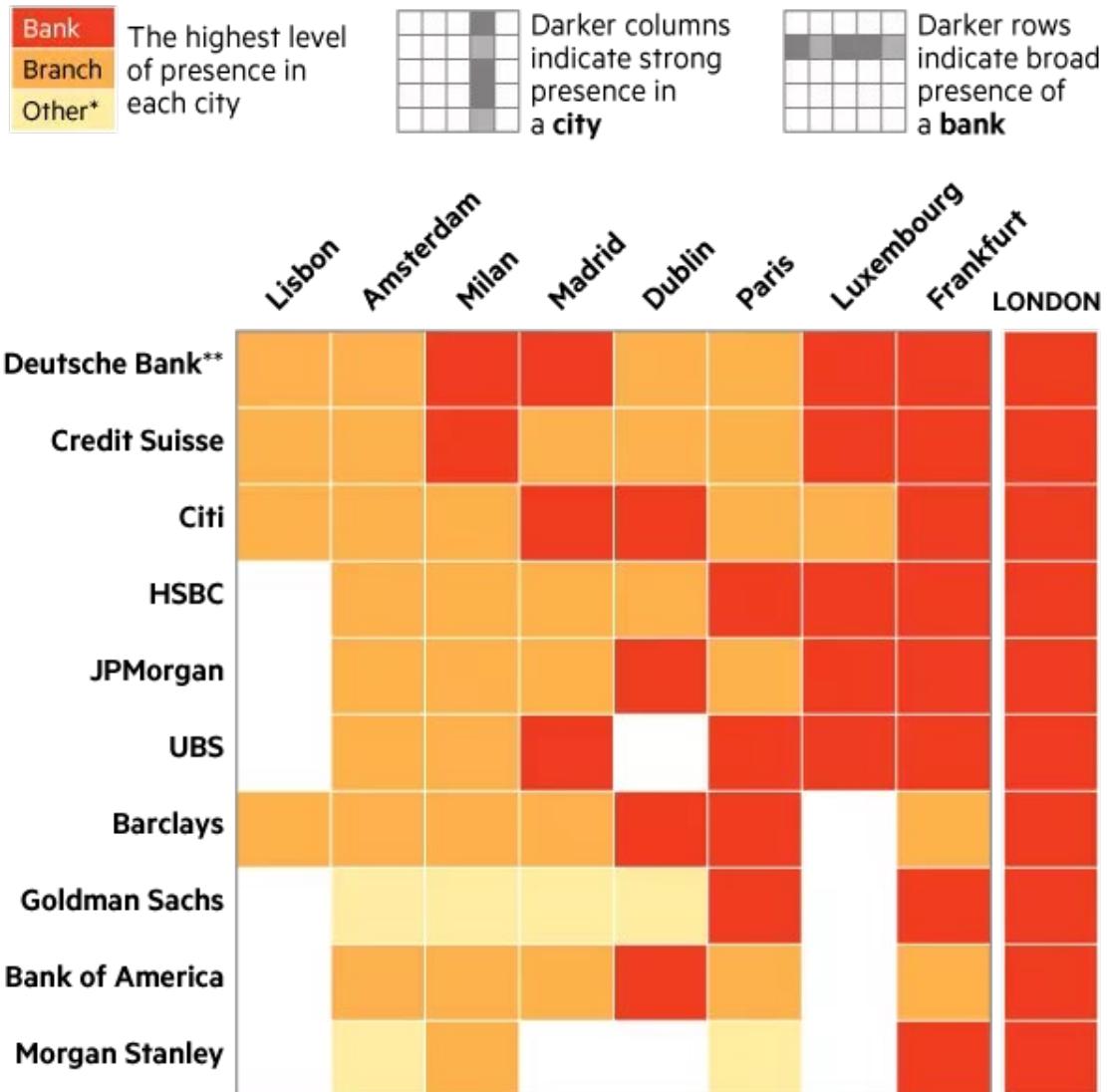


\* Broker dealer branches are included for Morgan Stanley and Goldman Sachs as they are a significant part of their European network

Deutsche Bank has a London subsidiary but its main entity is a branch

Source: FT research

## The Brexit banking matrix: The contenders lining up for London's crown



\* Broker dealer branches are included for Morgan Stanley and Goldman Sachs as they are a significant part of their European network

\*\* Deutsche Bank has a London subsidiary but its main entity is a branch

FT graphic Alan Smith, Laura Noonan Source: FT research

# Visualization Techniques

Lecture based on “Visualizing Geographical Data” course by Enrico Bertini

# Five Common Techniques

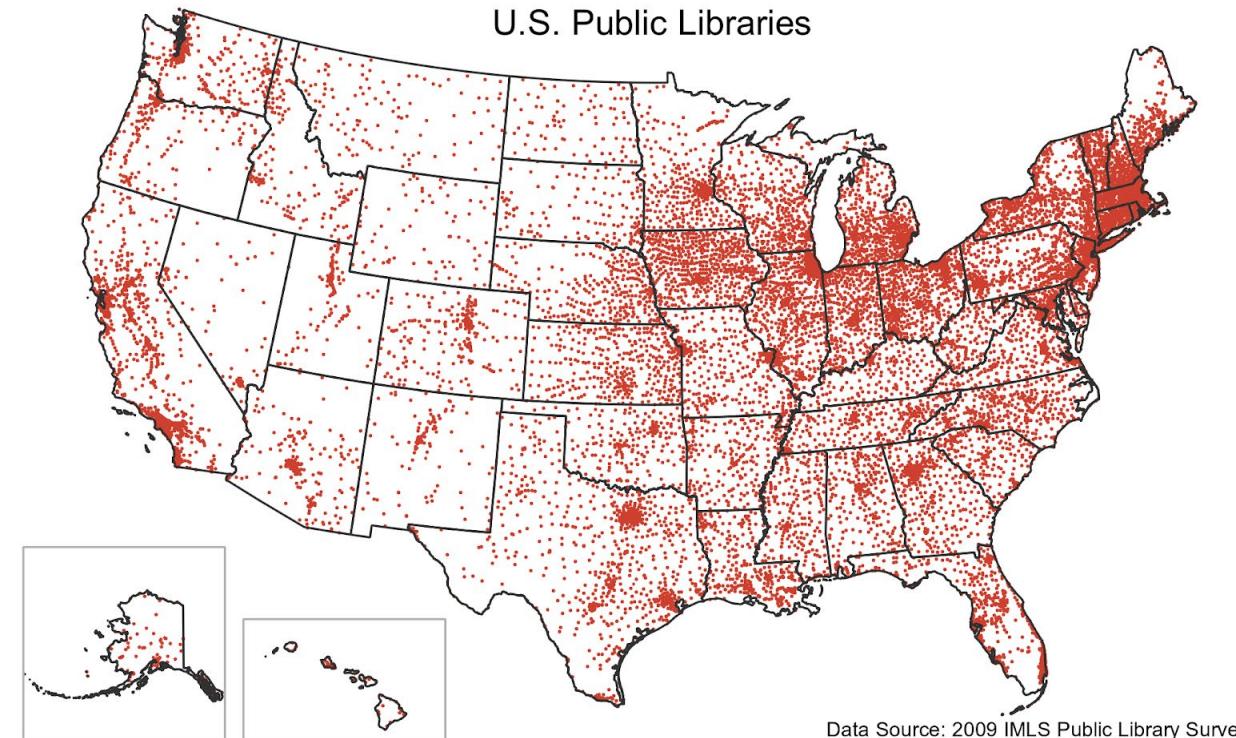
- Dot maps. (Distribution of objects.)
- Heat maps. (Distribution of values, continuous.)
- Binned maps. (Distribution of values, discrete/arbitrary.)
- Choropleth maps. (Distribution of values, discrete/geo.)
- Symbol maps. (Distribution of values, discrete/geo.)

# Dot Maps

To visualize the distribution of objects.

# Dot Maps

- Each item mapped to its location.
- Useful to depict **density** and **distribution** of events and geo-located objects.





## Home



Businesses can legally operate a sidewalk café if they hold a DCWP license or a temporary operating letter.

### Map Controls:

I want the map to show:

All Operating Cafés

Find a Business:

Enter a business name, address or ZIP code

Find a Location:

Enter an address or ZIP code

### Key:

#### SIDEWALK CAFÉ LOCATIONS

- Enclosed Sidewalk Café
  - Regular Unenclosed Sidewalk Café
  - Small Unenclosed Sidewalk Café
  - Combined Unenclosed Sidewalk Café
  - Inactive Sidewalk Café
- +  
+  
+  
+  
+

### File a Complaint:

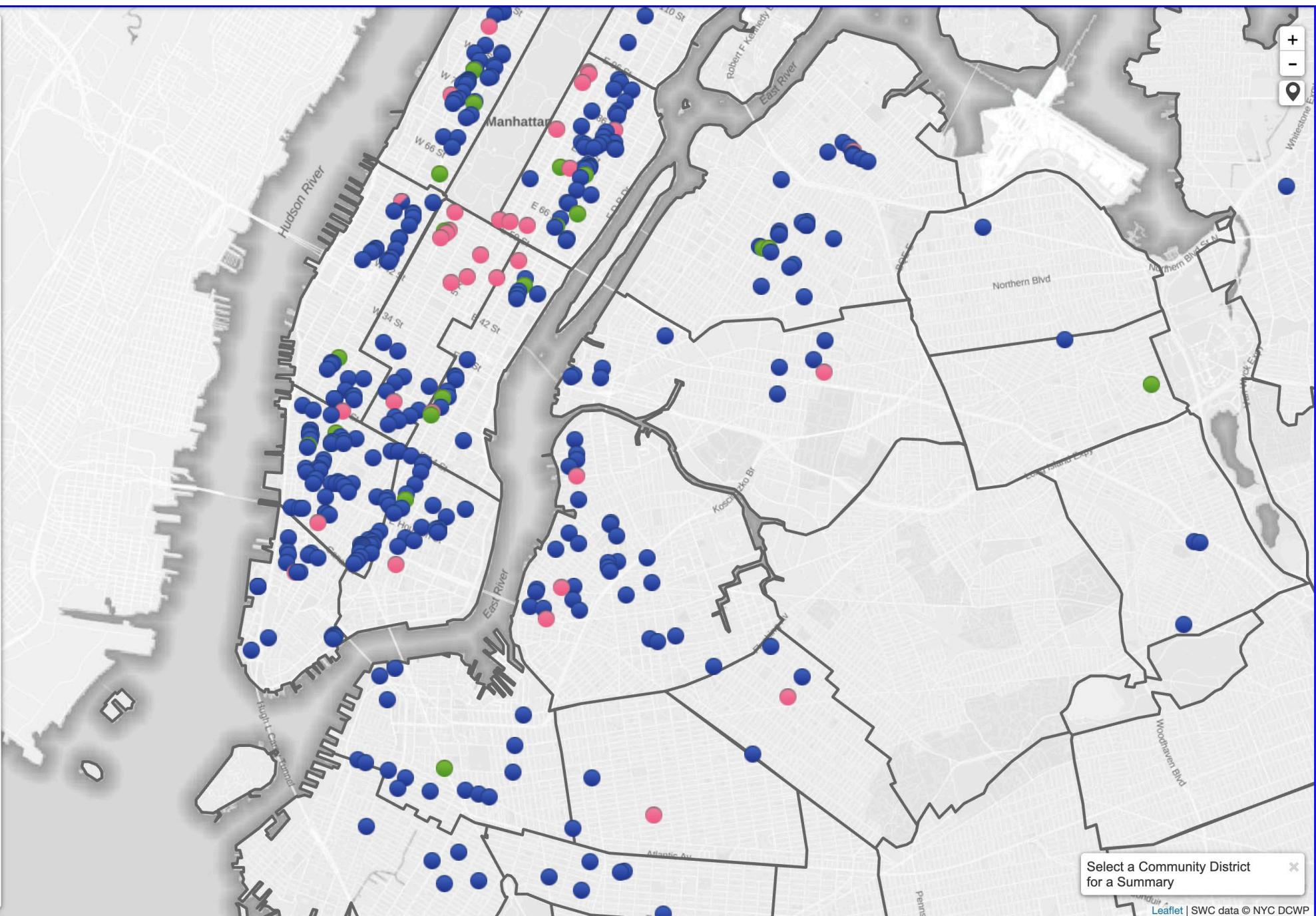
DCWP Website (account required) ↗

NYC 311 311 Online (no account required) ↗

### DISCLAIMER



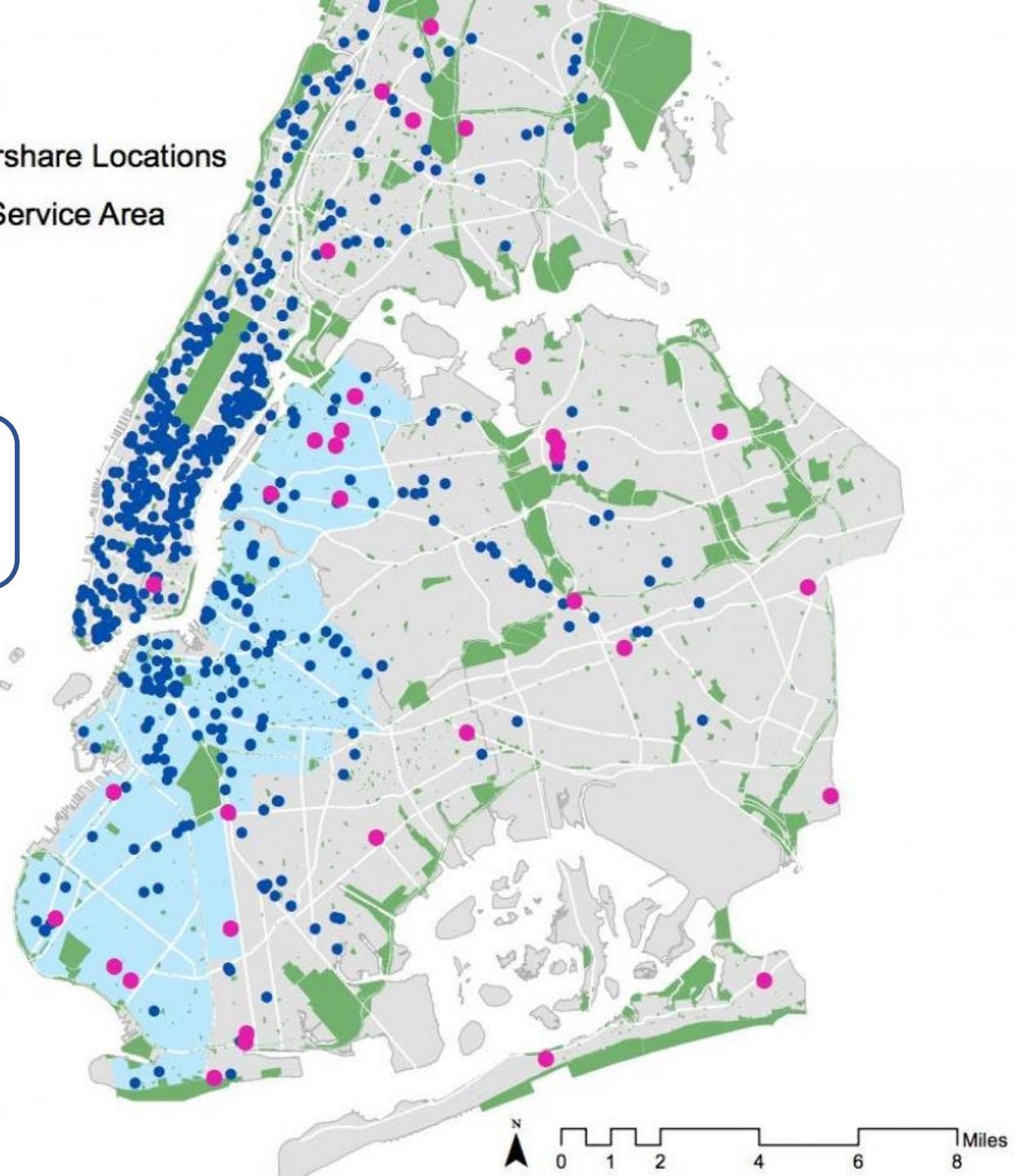
Built using Bootstrap, code licensed under [MIT](#).  
Sidebar icons licensed under [\(CC BY 3.0\)](#) by [Glyphicons](#).



- DOT Parking Facilities
- Zipcar & Enterprise Carshare Locations
- Car2Go & ReachNow Service Area



Categorical values can be encoded with color

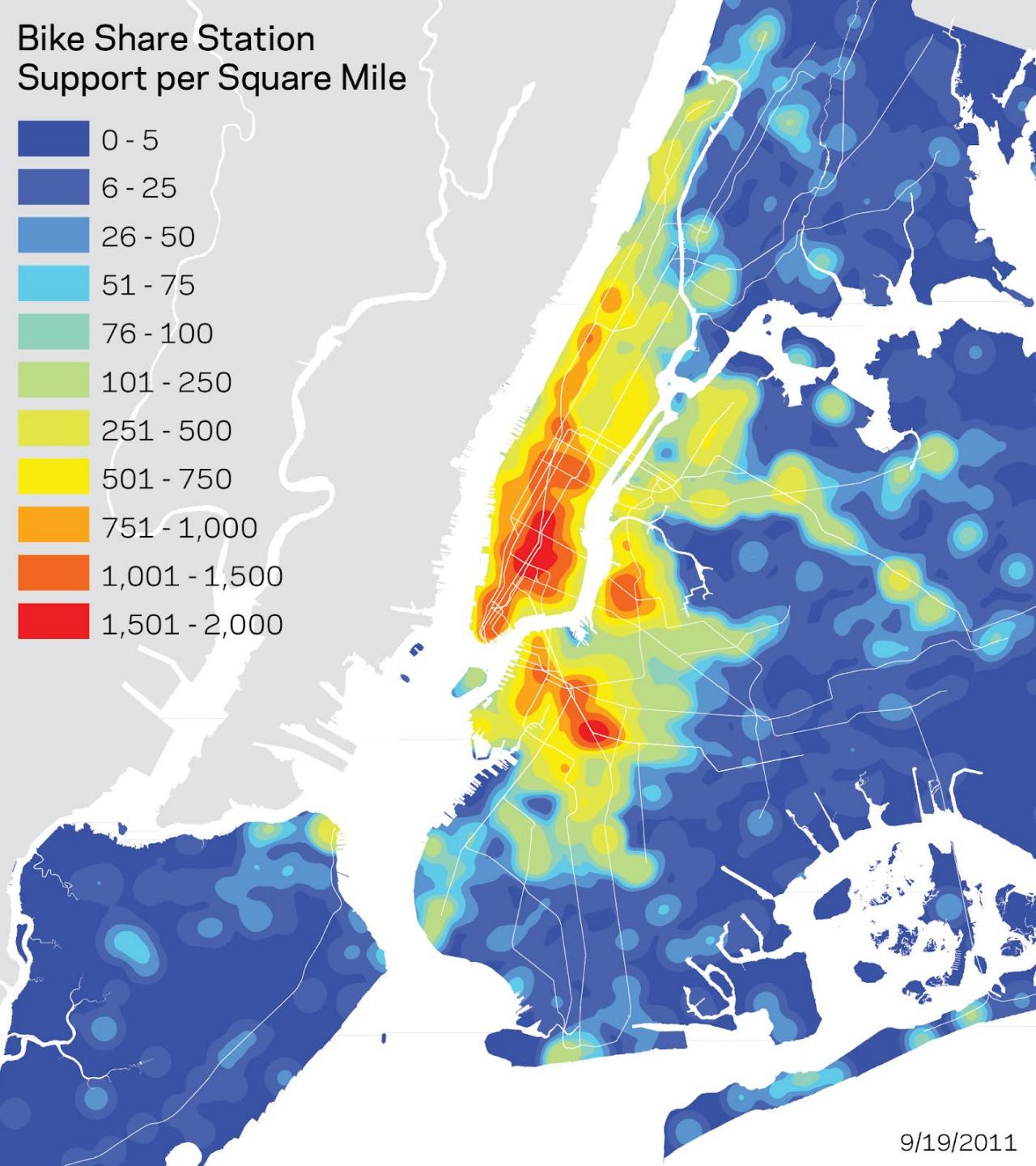
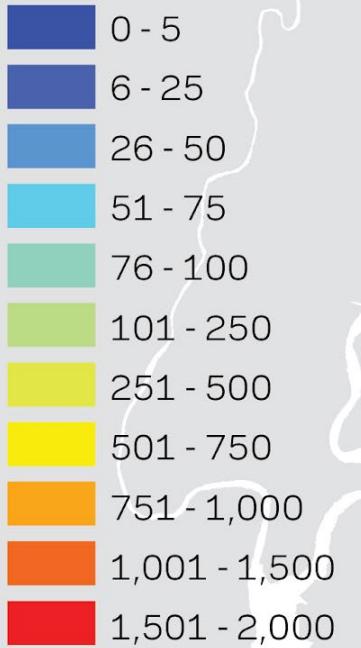


# Heat Maps

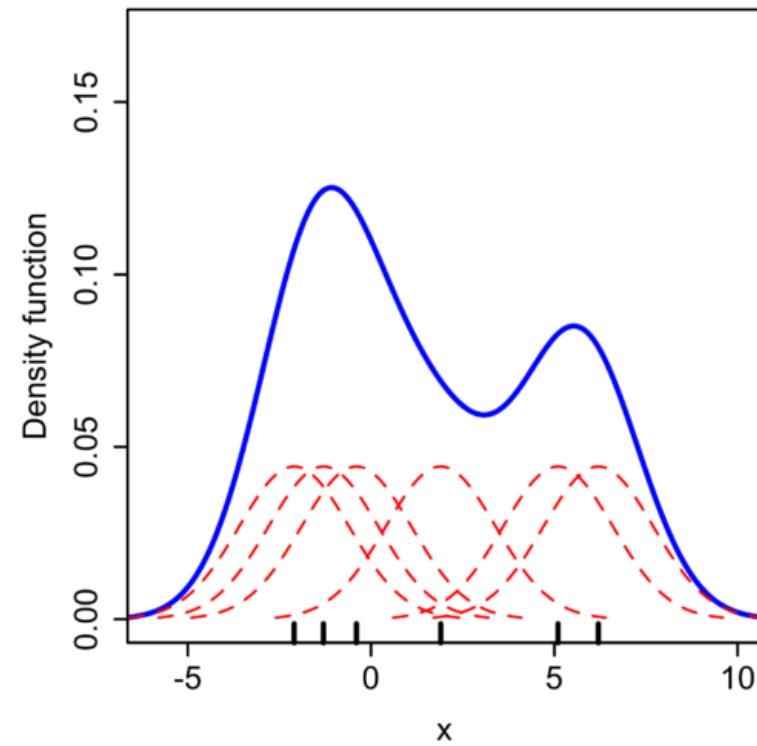
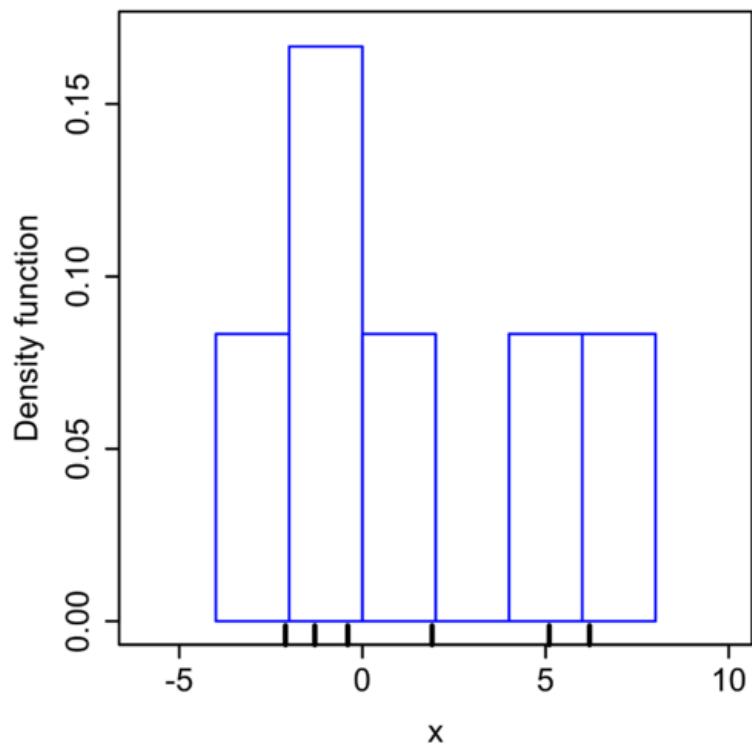
To visualize the continuous distribution of a quantity.

- Typically, used a “density estimation” method, which estimates a **continuous** density model from **discrete** data.

Bike Share Station Support per Square Mile

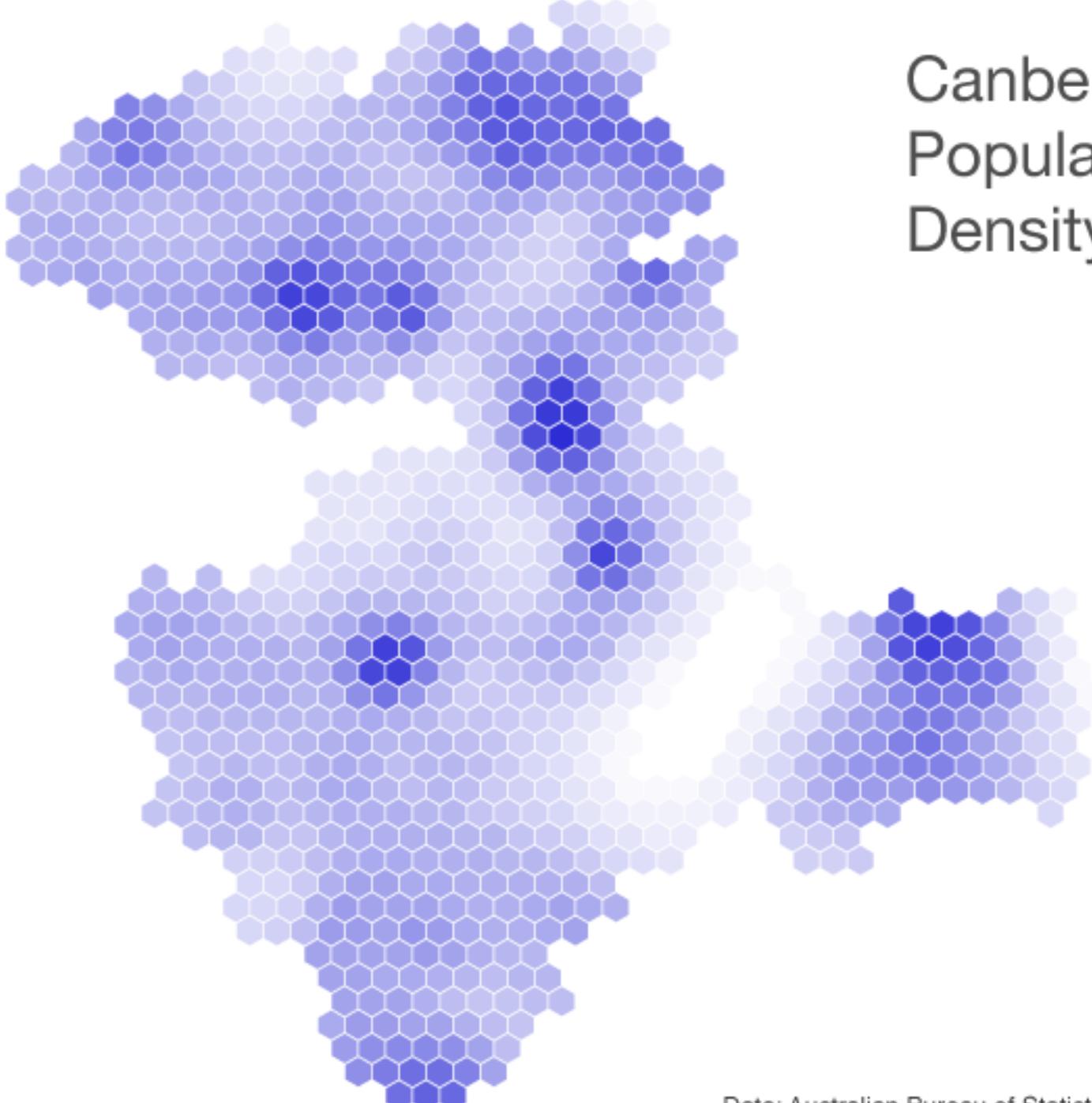


- Typically, used a “density estimation” method, which estimates a **continuous** density model from **discrete** data.



# Hexbin Maps

Visualize density differently.



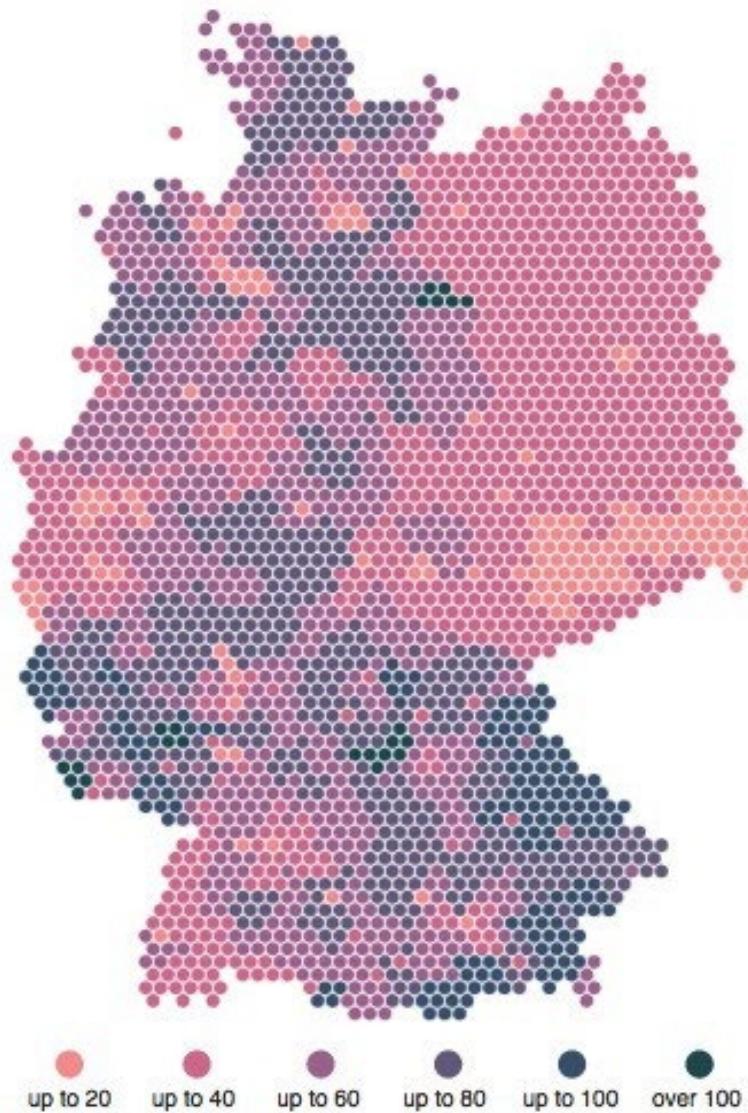
Canberra  
Population  
Density, 2011

Data: Australian Bureau of Statistics, 2011 Census.



## Gun Ownership

Number of legal handguns per 1,000 inhabitants, 2013

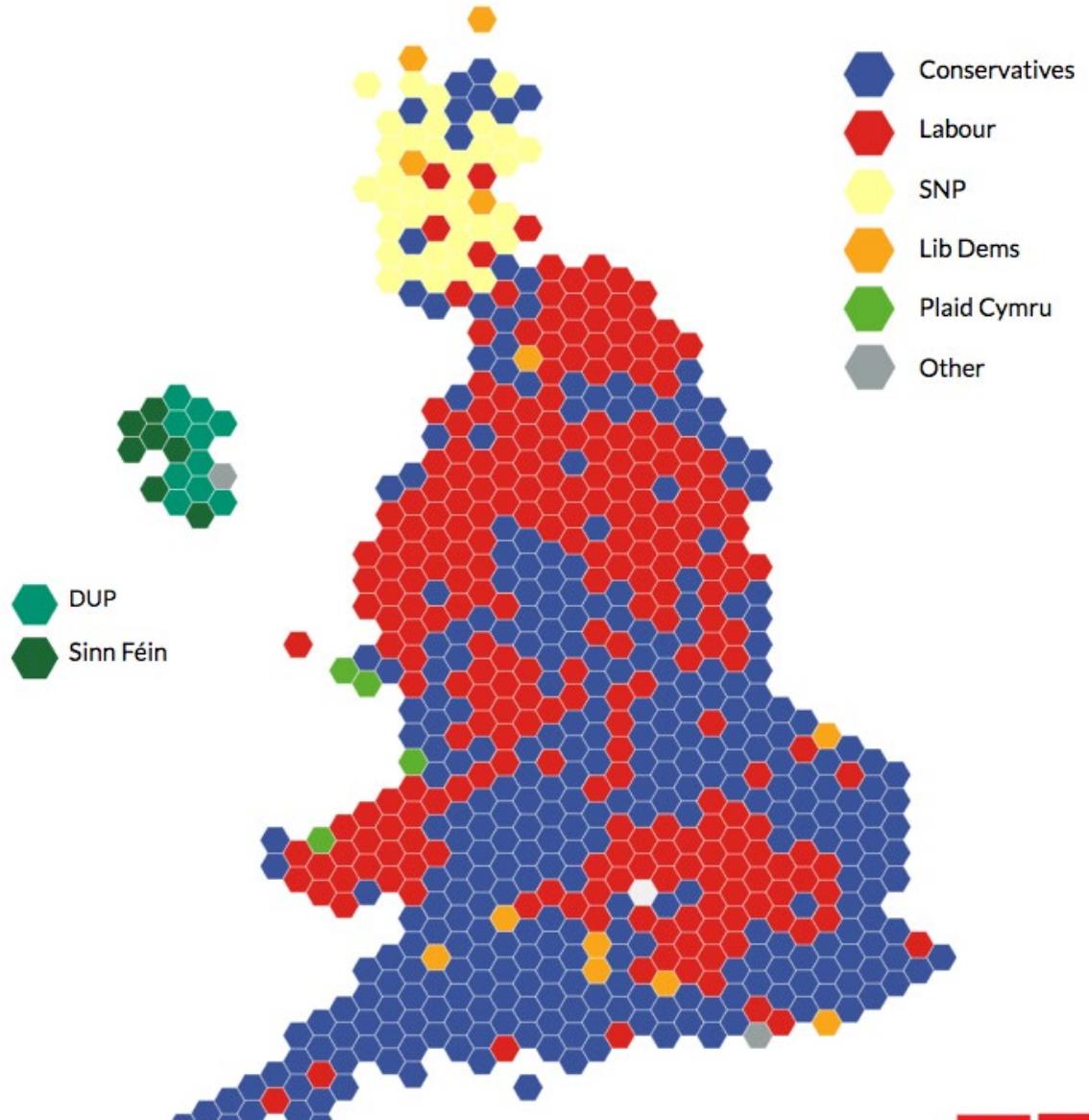




# The results

## Categorical Information

Source: <https://www.ibtimes.co.uk/john-mcdonnell-pro-brexit-labour-will-take-uk-out-eus-single-market-1625757>



# Choropleth Maps

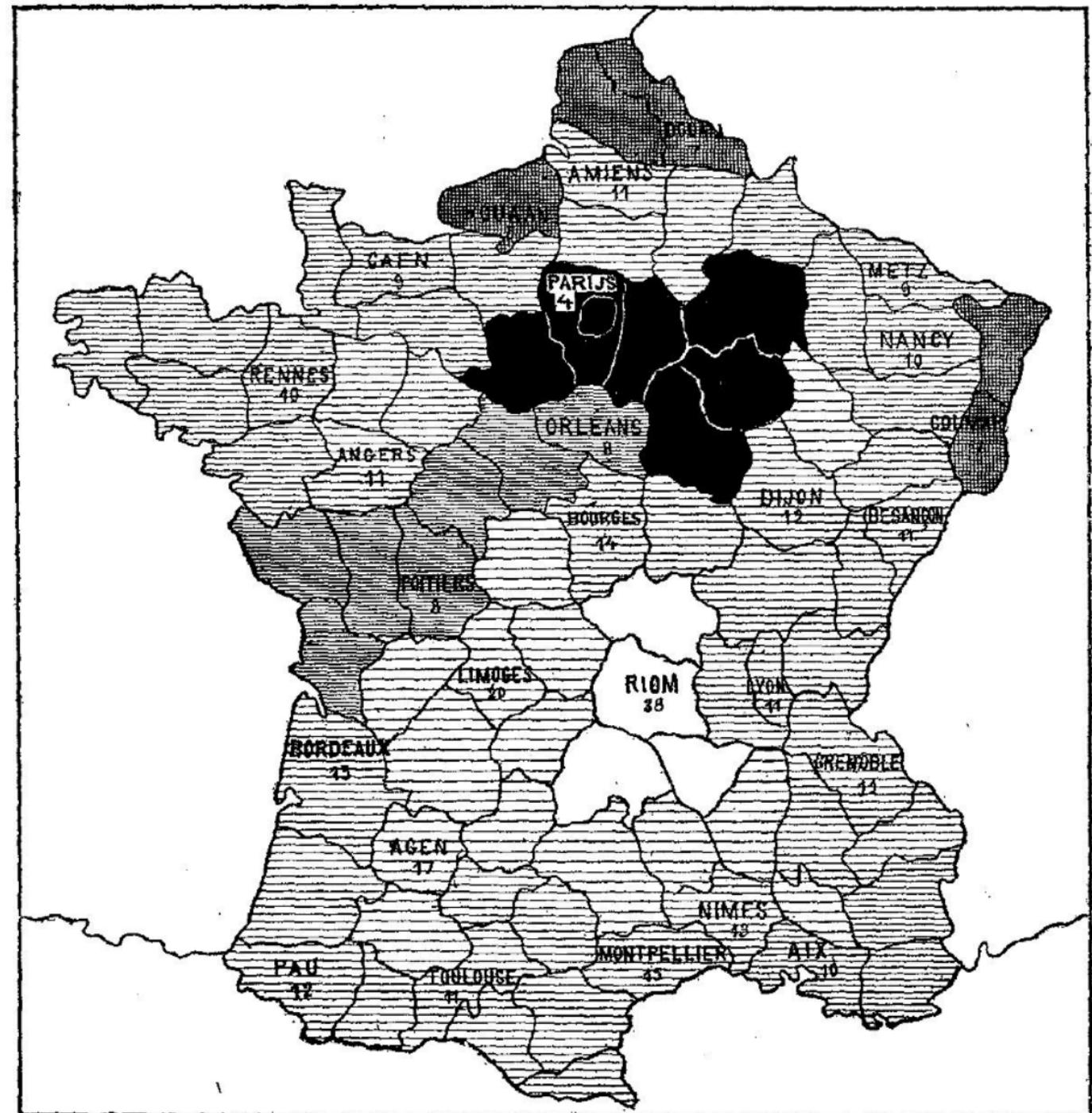
Convey rates across regions.

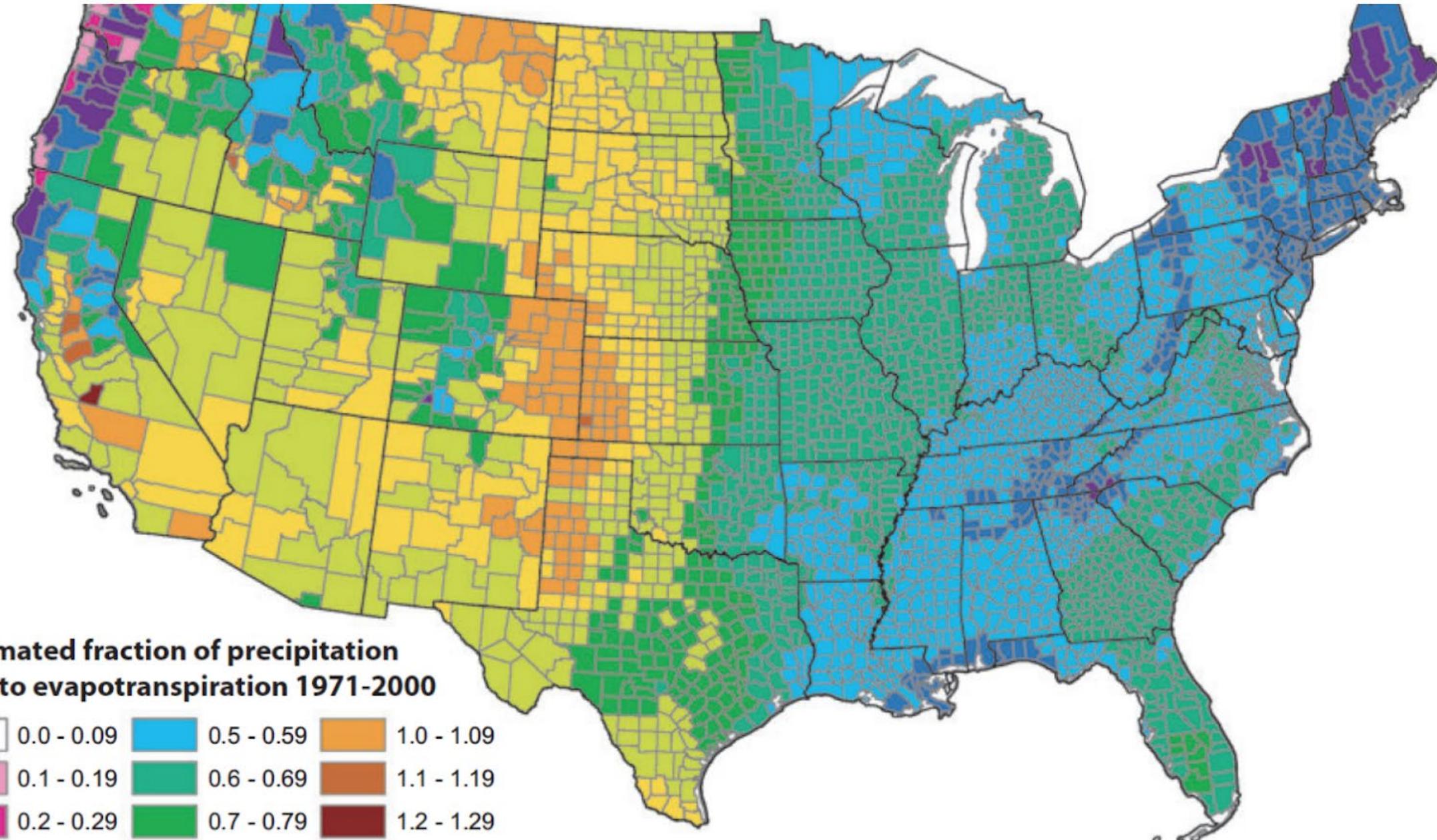
# 1826(?) Literacy in France

By Pierre Charles Dupin

Source:

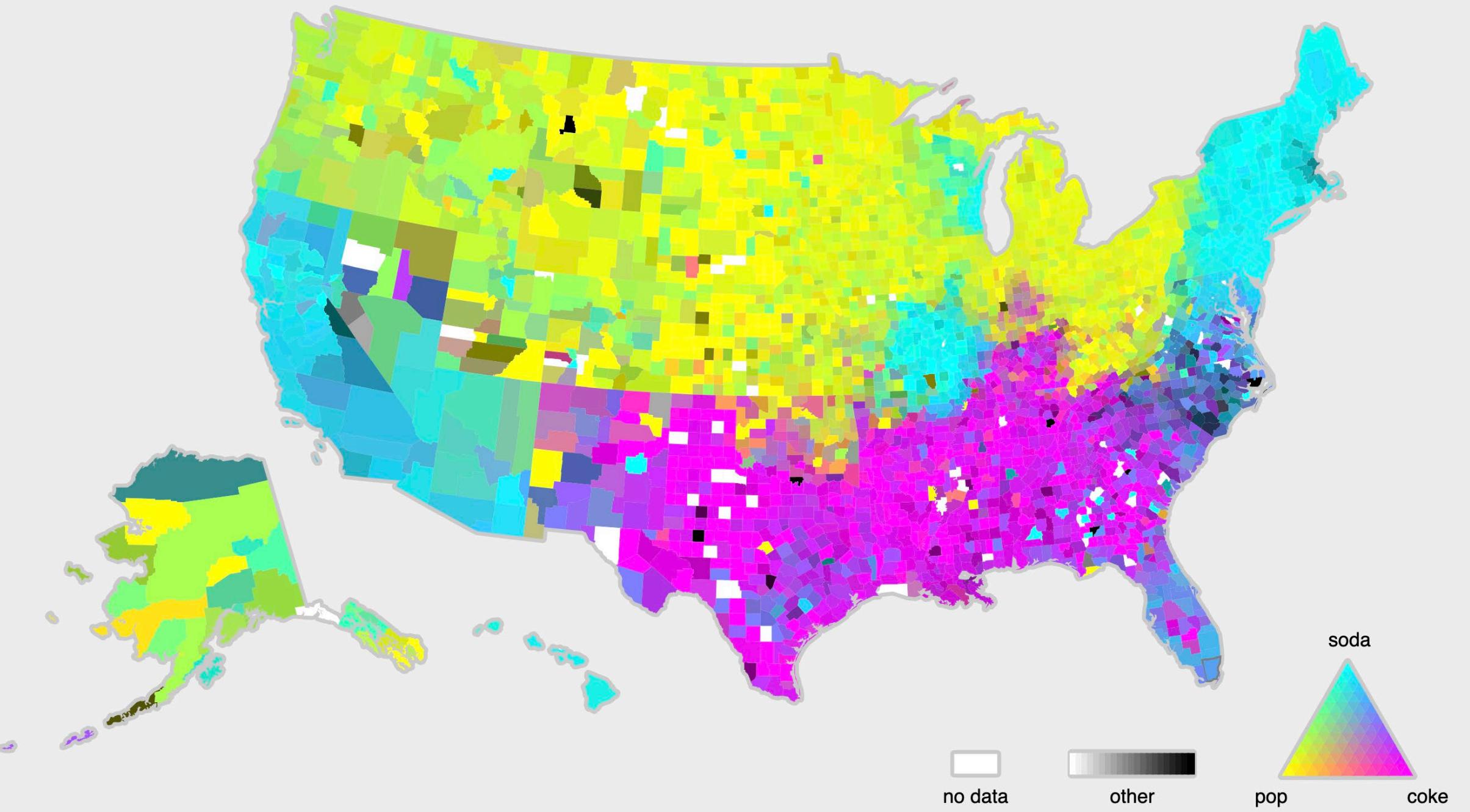
<https://courses.cs.washington.edu/courses/cse512/21sp/lectures/CSE512-Maps.pdf>

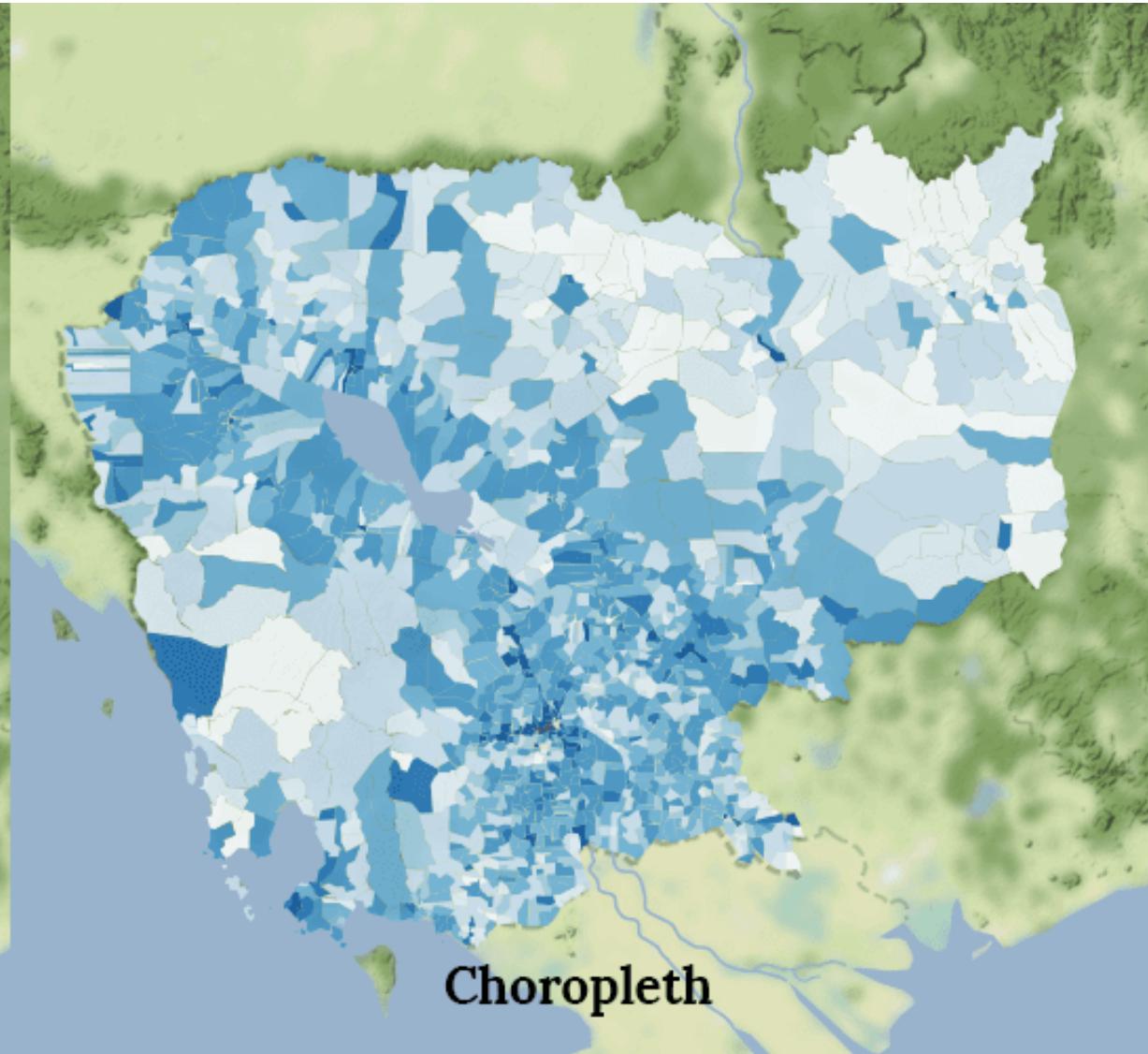
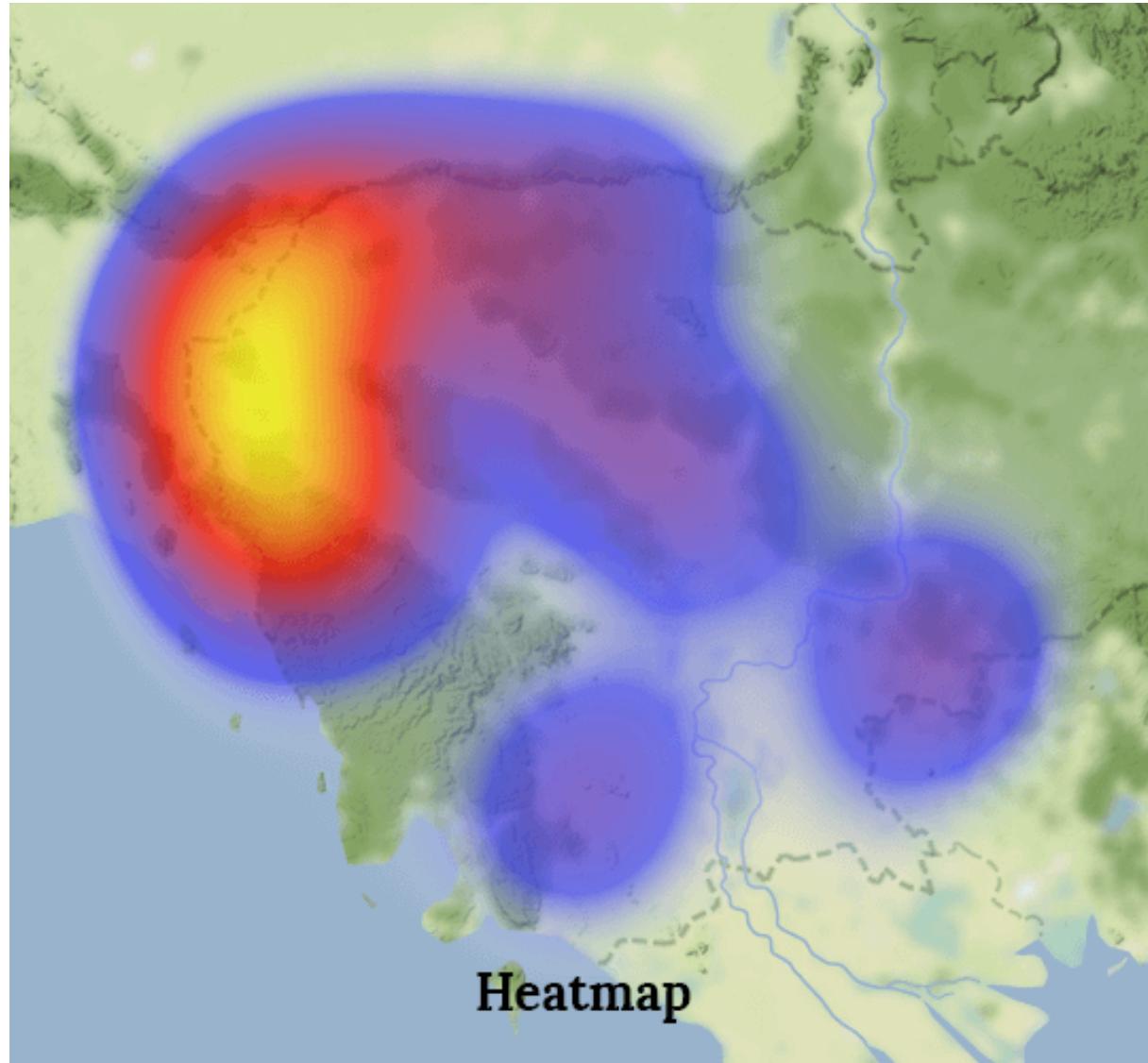




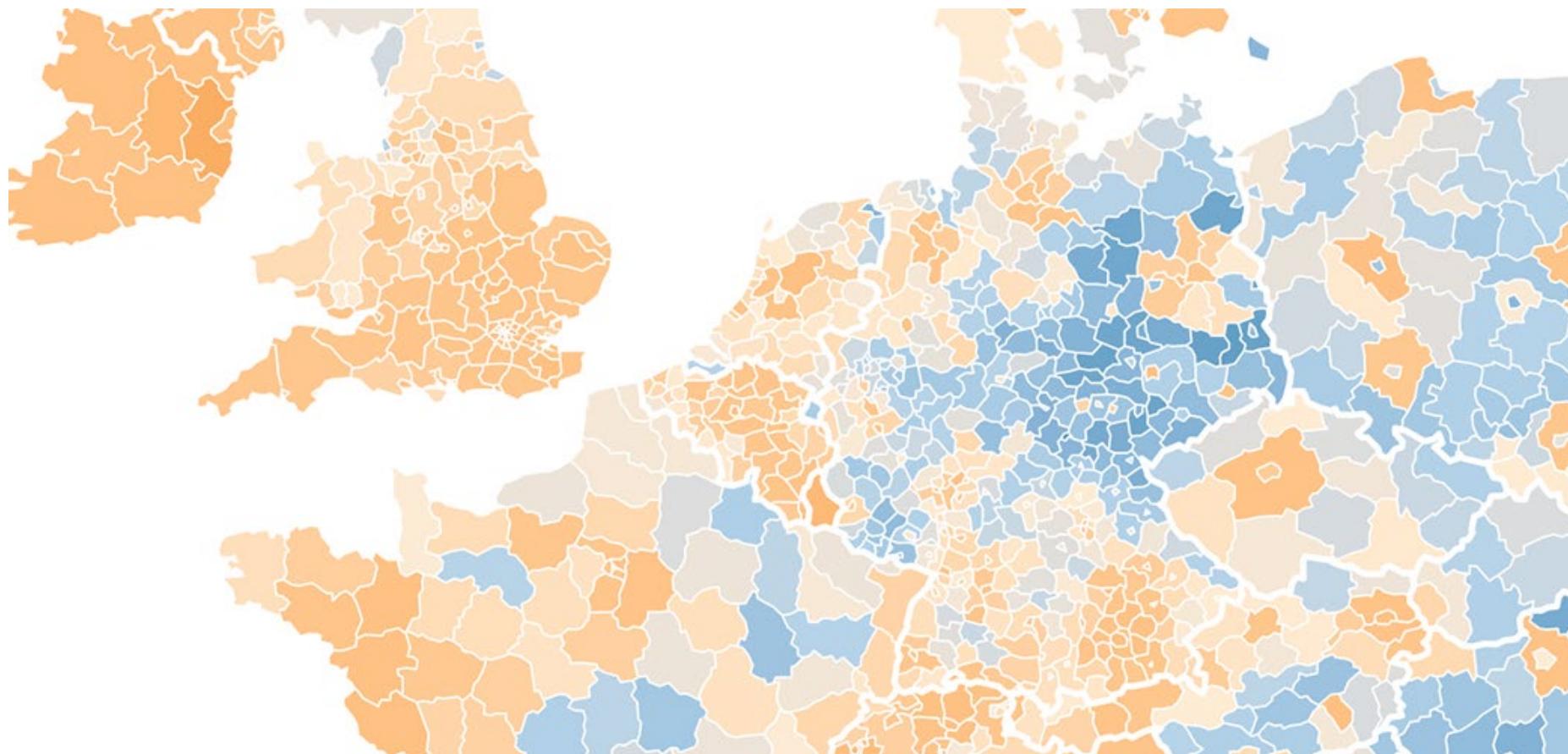
**Estimated fraction of precipitation  
lost to evapotranspiration 1971-2000**

[white]	0.0 - 0.09	[light blue]	0.5 - 0.59	[orange]	1.0 - 1.09
[pink]	0.1 - 0.19	[green]	0.6 - 0.69	[brown]	1.1 - 1.19
[magenta]	0.2 - 0.29	[dark green]	0.7 - 0.79	[dark brown]	1.2 - 1.29
[purple]	0.3 - 0.39	[yellow-green]	0.8 - 0.89		
[blue]	0.4 - 0.49	[yellow]	0.9 - 0.99		



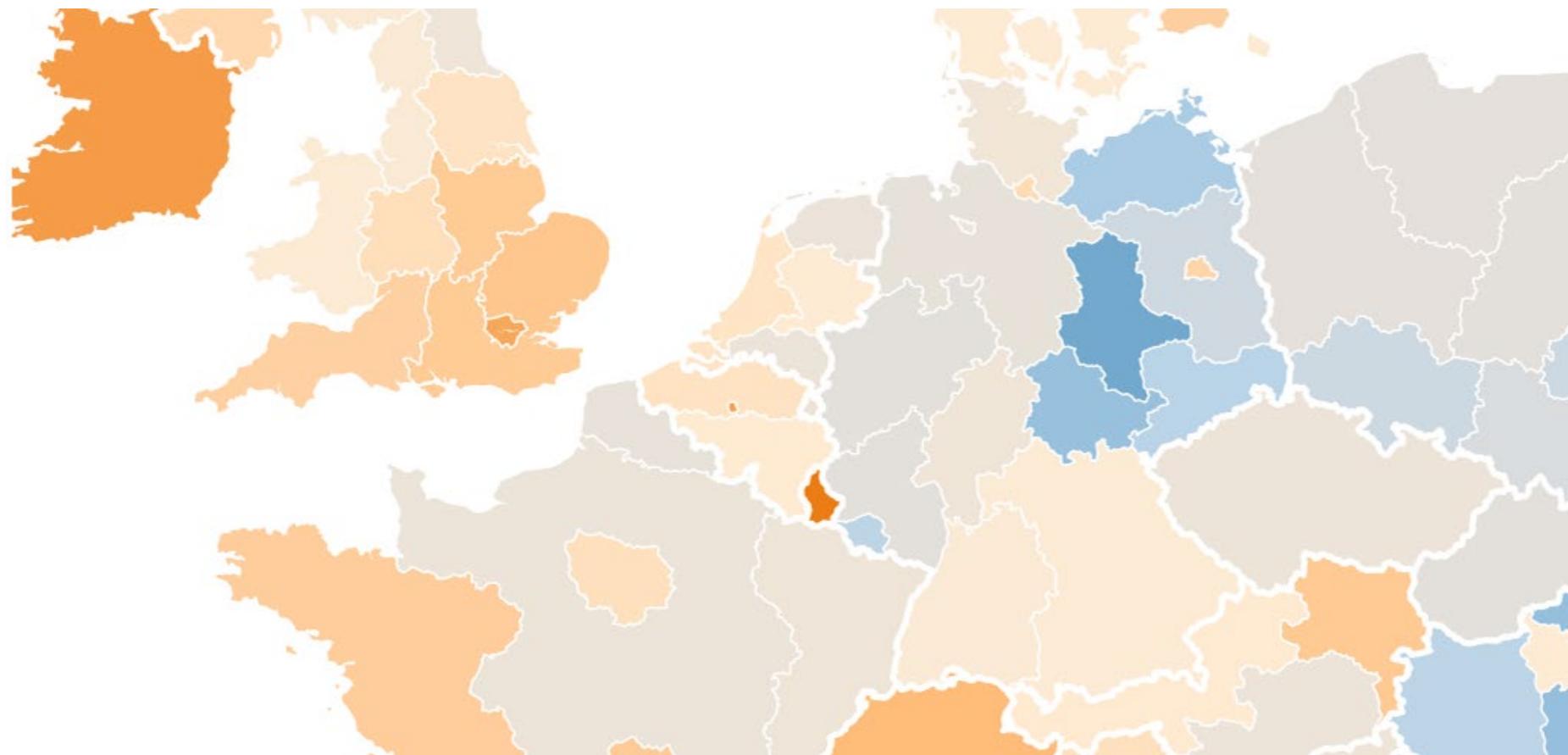


# Level of Details



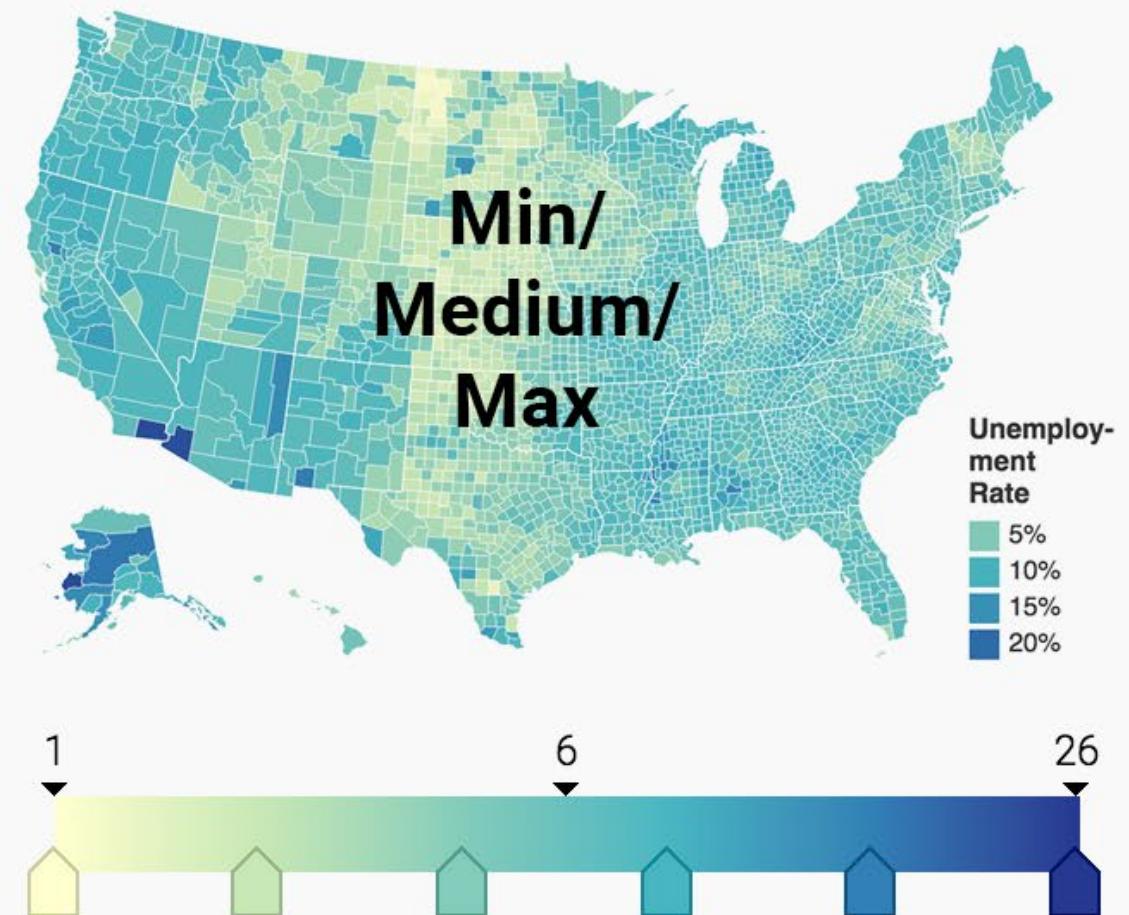
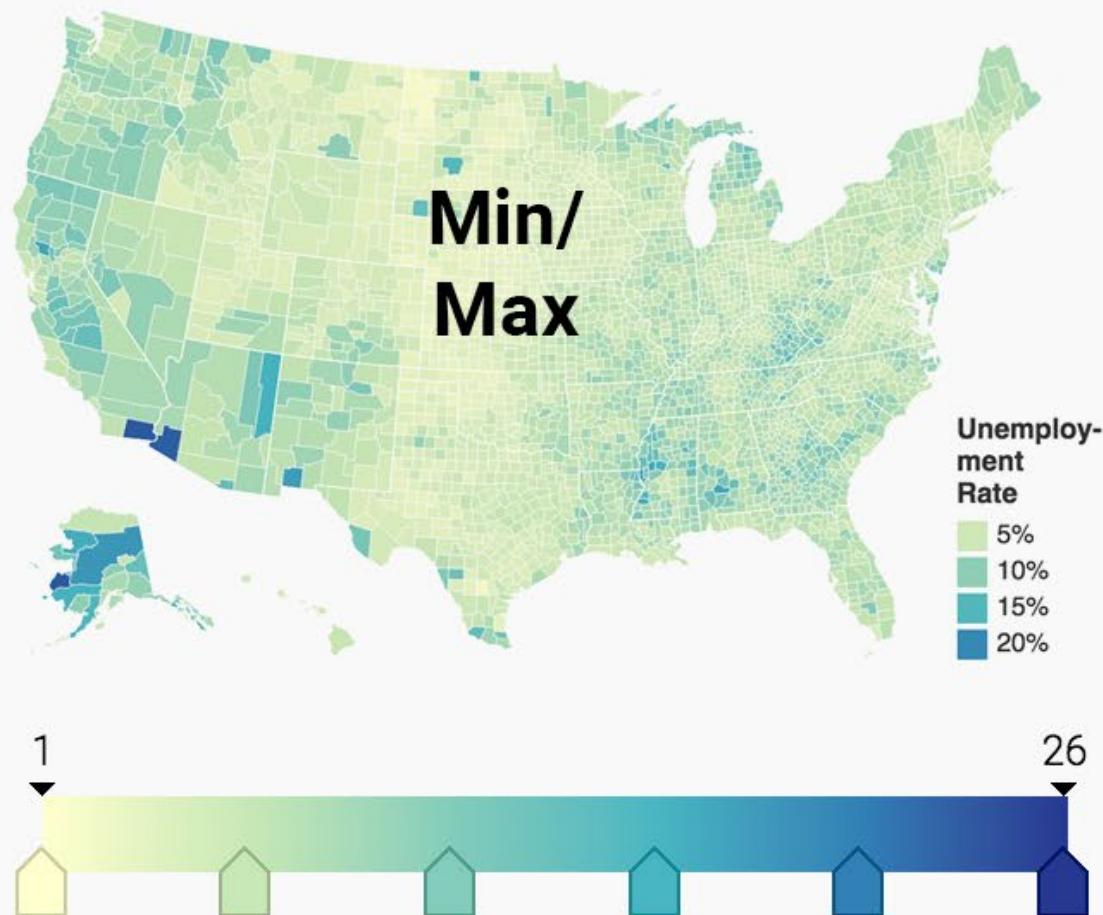
Source: <https://blog.datawrapper.de/weekly-chart-europegrowth/>

# Level of Details



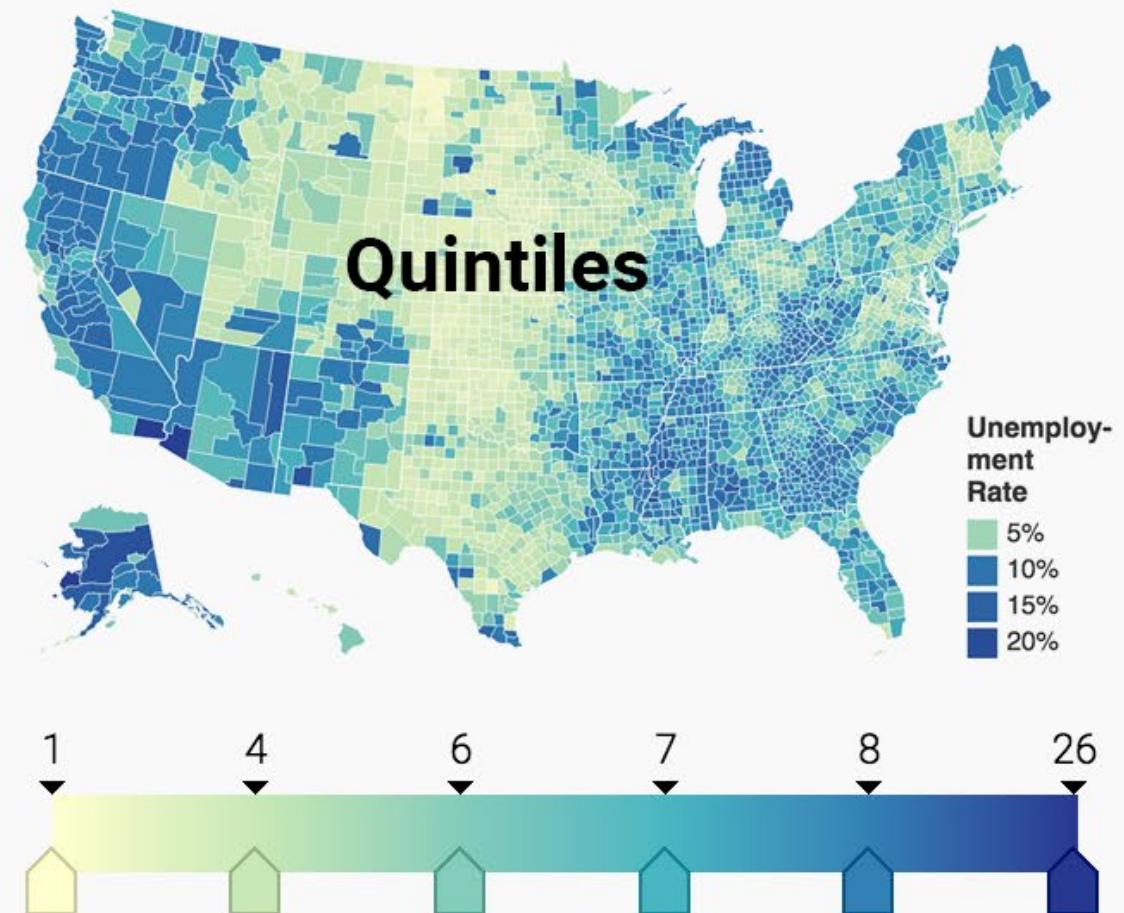
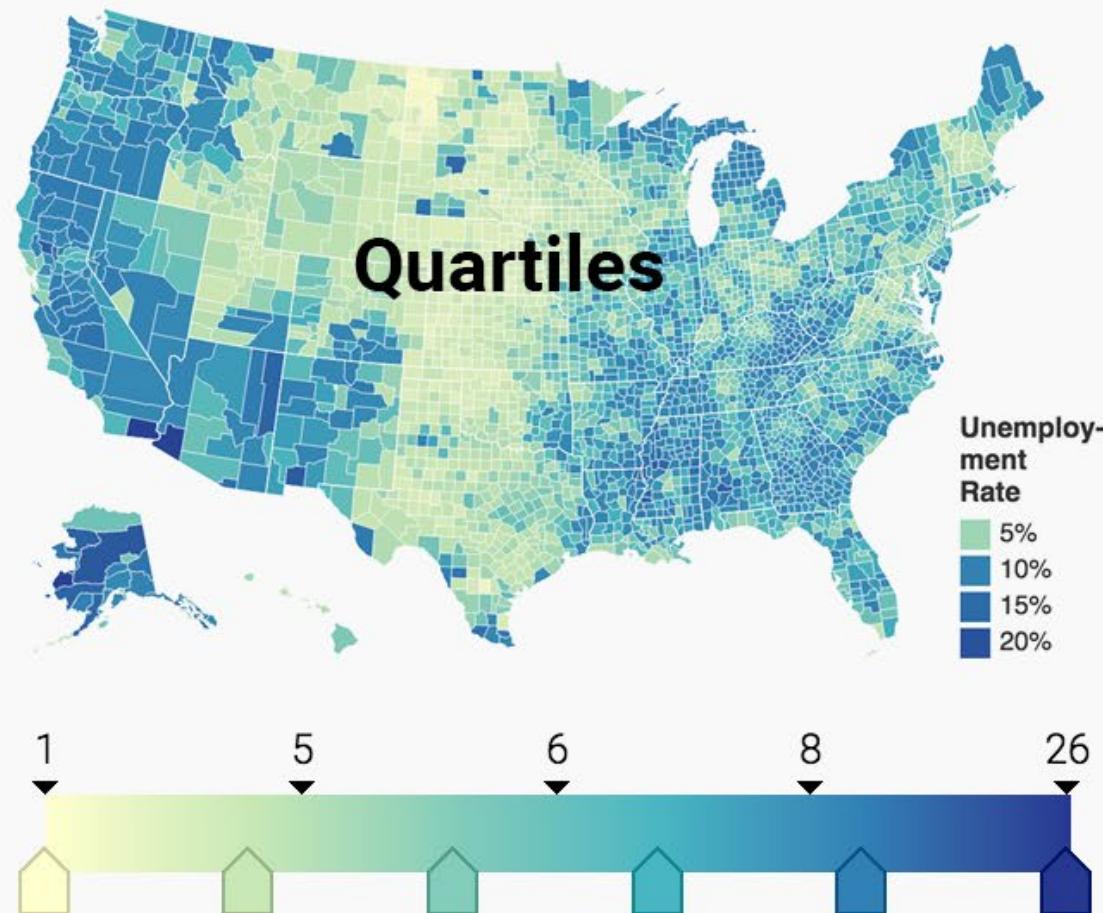
Source: <https://blog.datawrapper.de/weekly-chart-europegrowth/>

# Color matters



Source: <https://academy.datawrapper.de/article/117-color-palette-for-your-map>

# Color matters

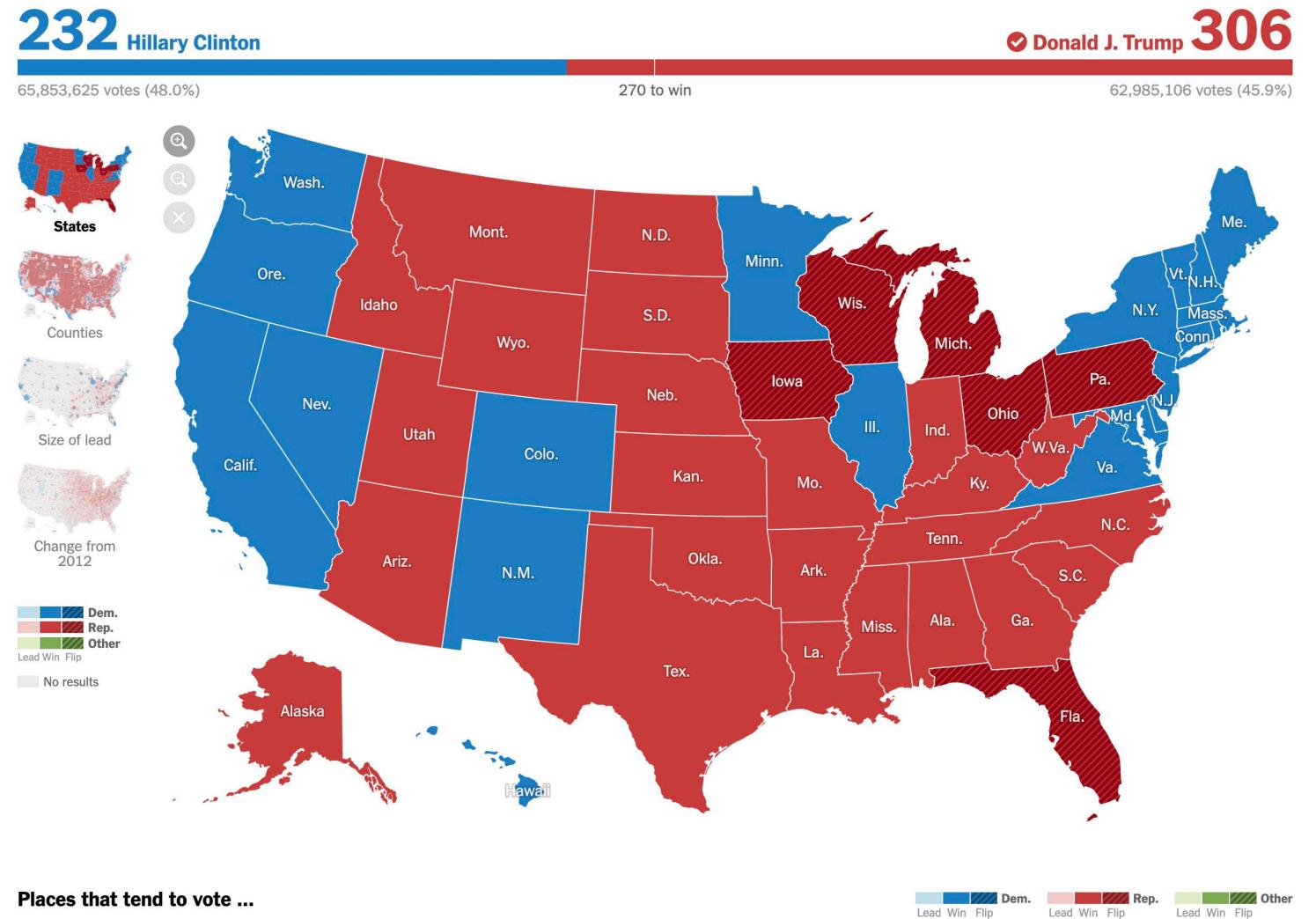


Source: <https://academy.datawrapper.de/article/117-color-palette-for-your-map>

# Categorical information

Note that no quantity information is associated with color.

Source:  
<https://www.nytimes.com>  
president

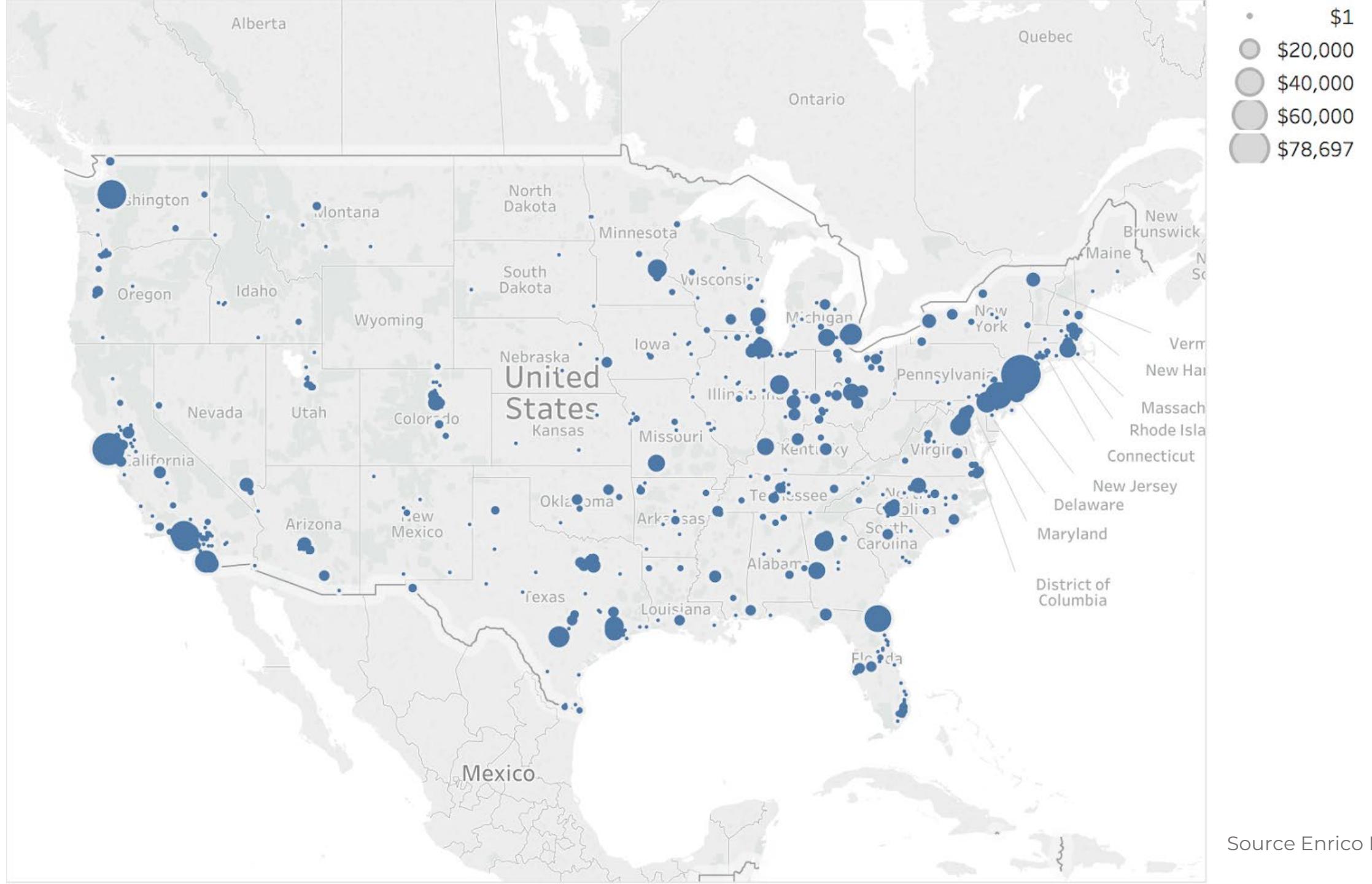


# Focus on the foreground



# Graduated Symbol Maps

Using size of symbol to represent quantity.



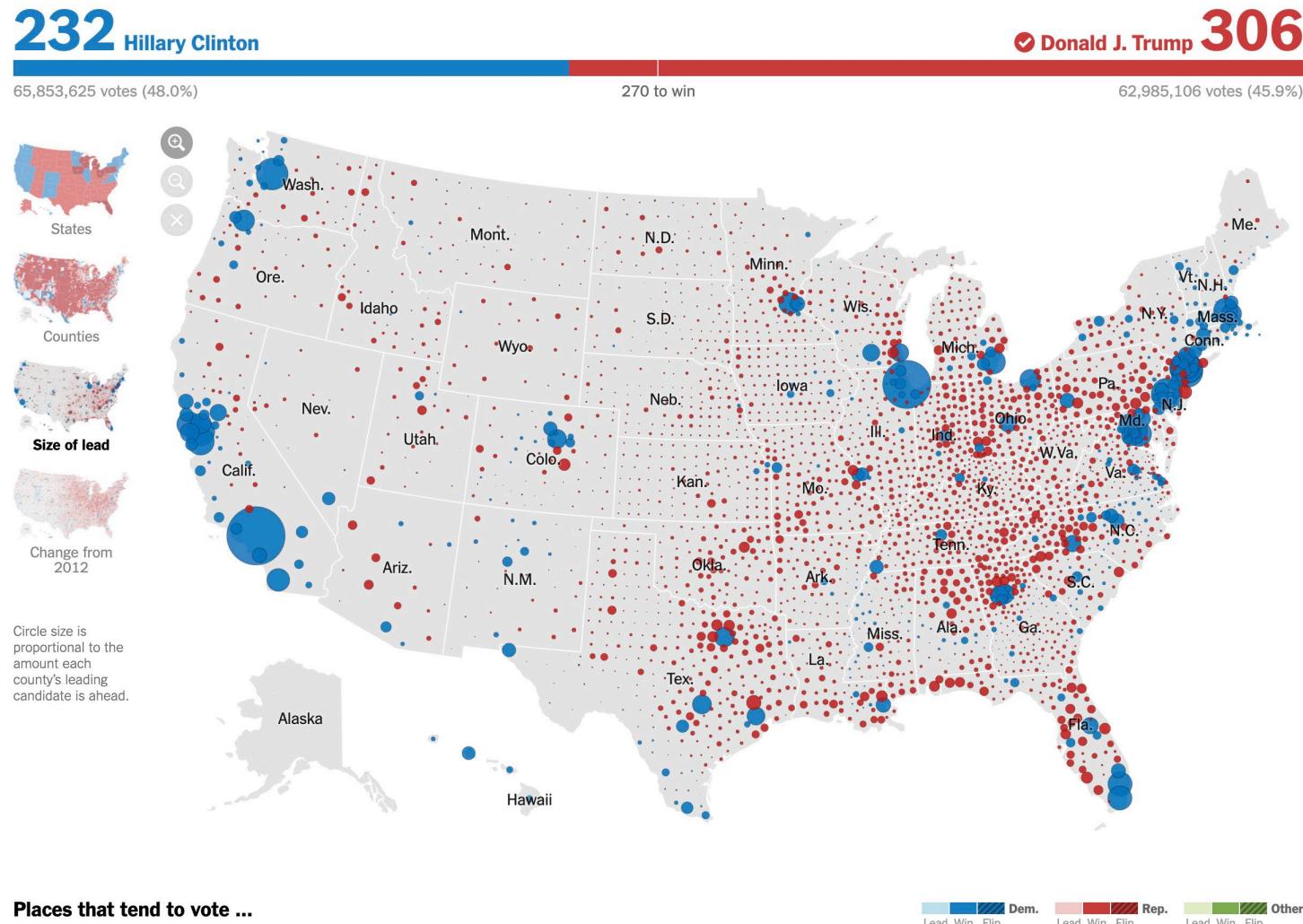
Source Enrico Bertini

## Graduated Symbol Maps

- Does not depend on regions' size/shape.
- Easier to use multiple channels.
- Do not solve the spatial density problem.

Source:

<https://www.nytimes.com/elections/2016/results/president>



# Immigration Explorer

Select a foreign-born group to see how they settled across the United States.

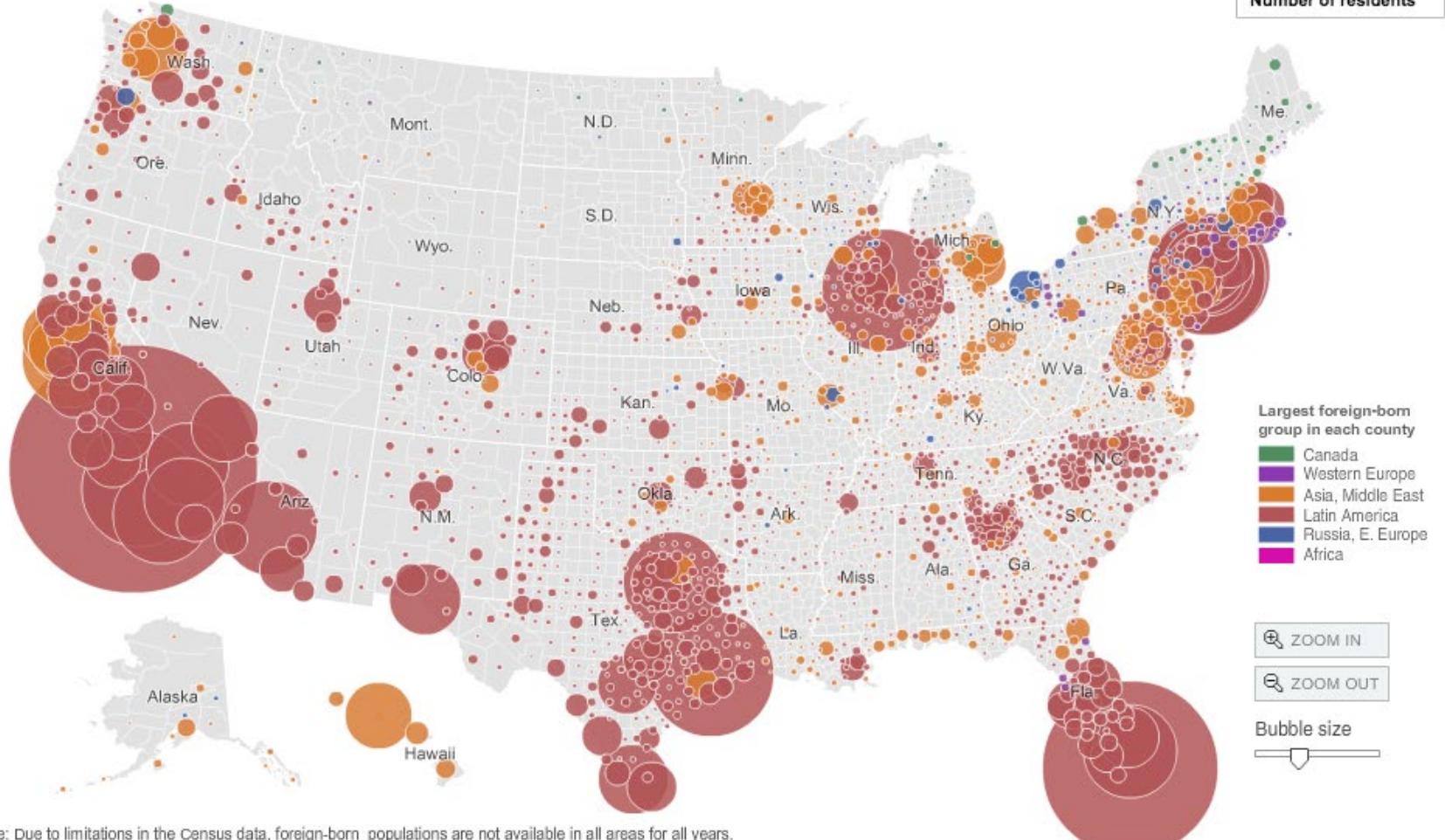
All Countries

1880 '90 1900 '10 1920 '30 1940 '50 1960 '70 1980 '90 2000

SHOW

Percent of population

Number of residents

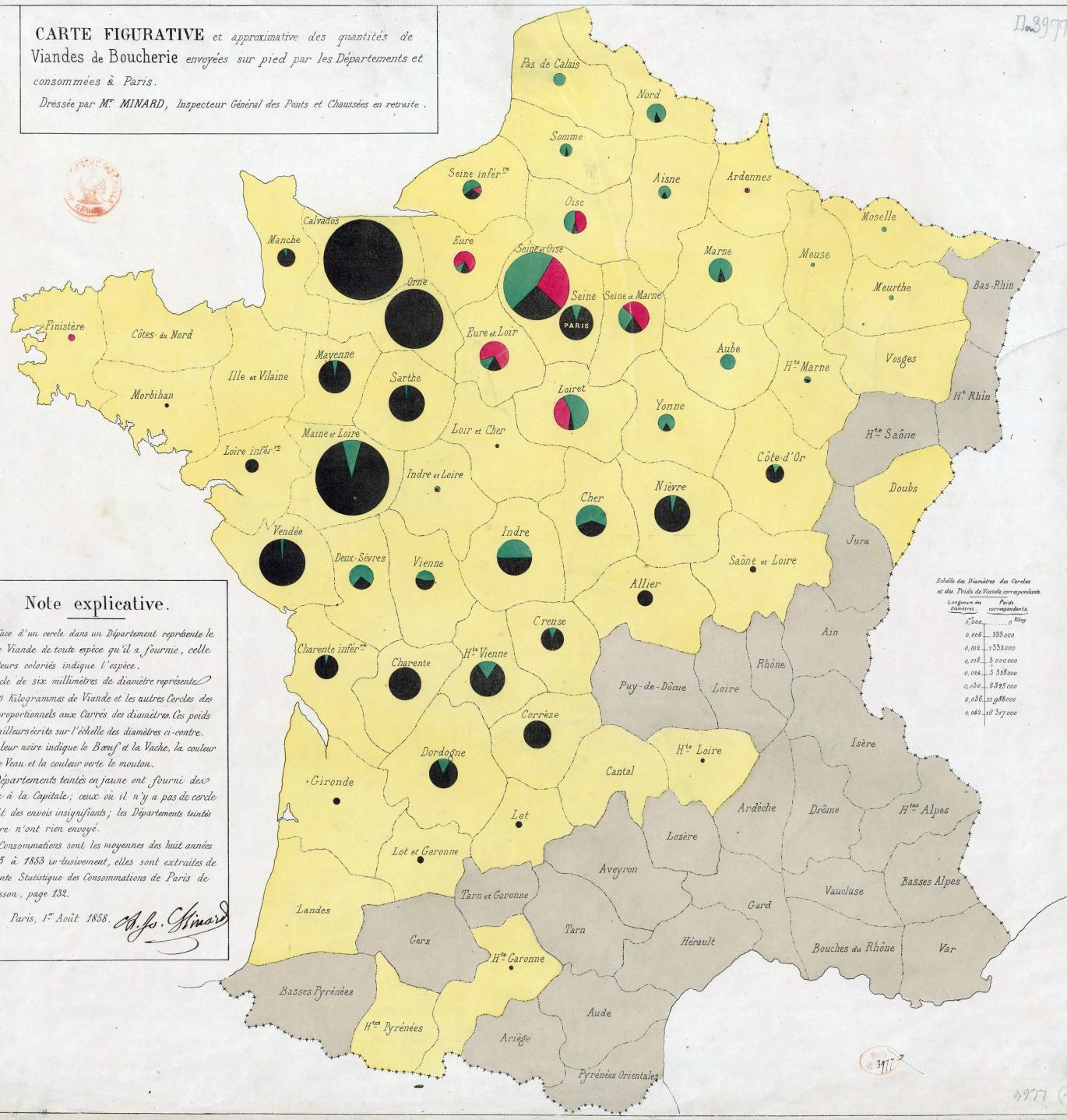


Note: Due to limitations in the Census data, foreign-born populations are not available in all areas for all years.

Sources: Social Explorer, [www.socialexplorer.com](http://www.socialexplorer.com); Minnesota Population Center; U.S. Census Bureau

Matthew Bloch and Robert Gebeloff/The New York Times

CARTE FIGURATIVE et approximative des quantités de  
Viandes de Boucherie envoyées sur pied par les Départements et  
consommées à Paris.  
Dressée par M<sup>e</sup> MINARD, Inspecteur Général des Ponts et Chaussées en retraite.



## Complex Symbols

Cattle from all around France  
By Charles Joseph Minard  
(1858)

Source:

[https://en.wikipedia.org/wiki/Charles\\_Joseph\\_Minard#/media/File:Minard-carte-viande-1858.png](https://en.wikipedia.org/wiki/Charles_Joseph_Minard#/media/File:Minard-carte-viande-1858.png)

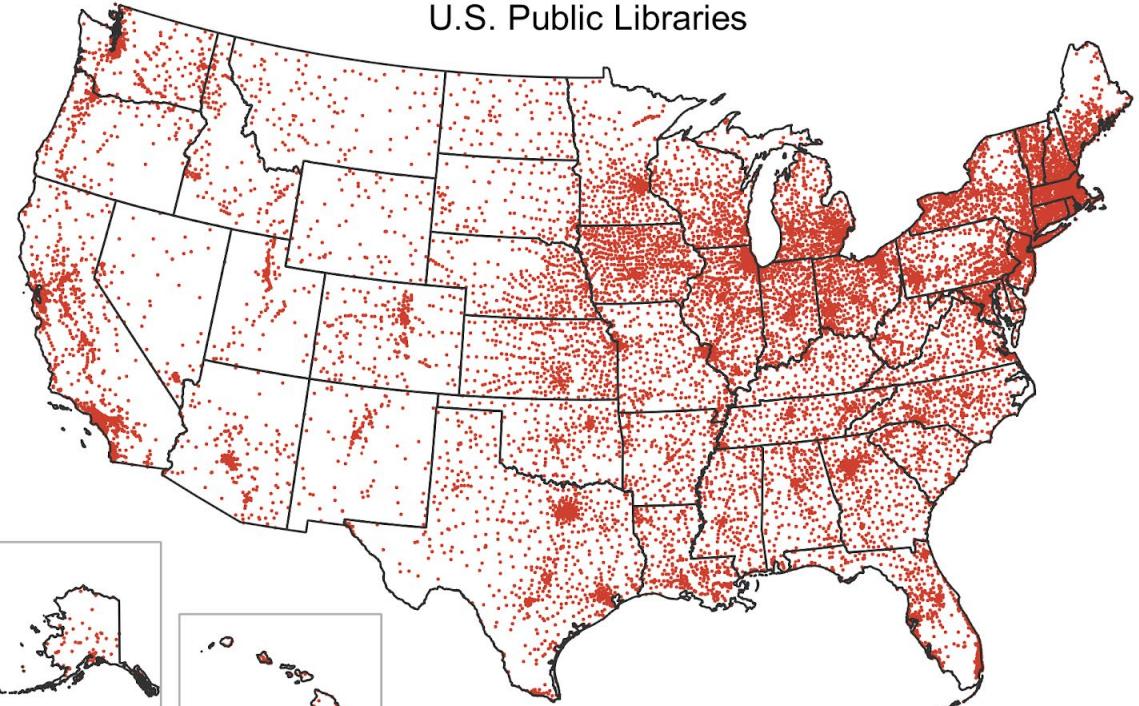
# Summary

- Dot maps. (Distribution of objects.) Objects
- Heat maps. (Distribution of values, continuous.) Densities
- Binned maps. (Distribution of values, discrete/arbitrary.)
- Choropleth maps. (Distribution of values, discrete/geo.) Values
- Symbol maps. (Distribution of values, discrete/geo.)

# Common Issues with Maps

# Base Rate Bias

U.S. Public Libraries



Data Source: 2009 IMLS Public Library Survey

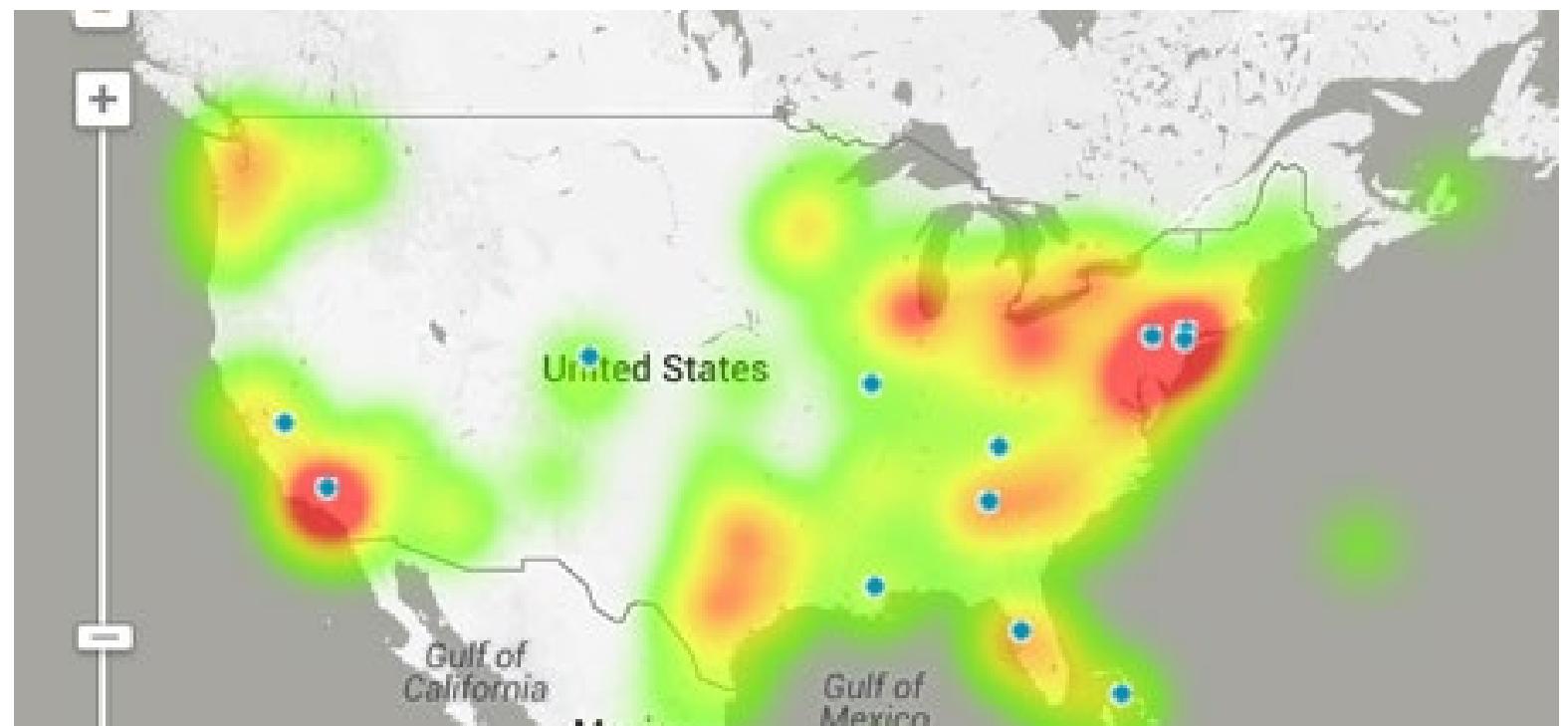
US Population Map ([www.populationlabs.com](http://www.populationlabs.com))



## True or False?

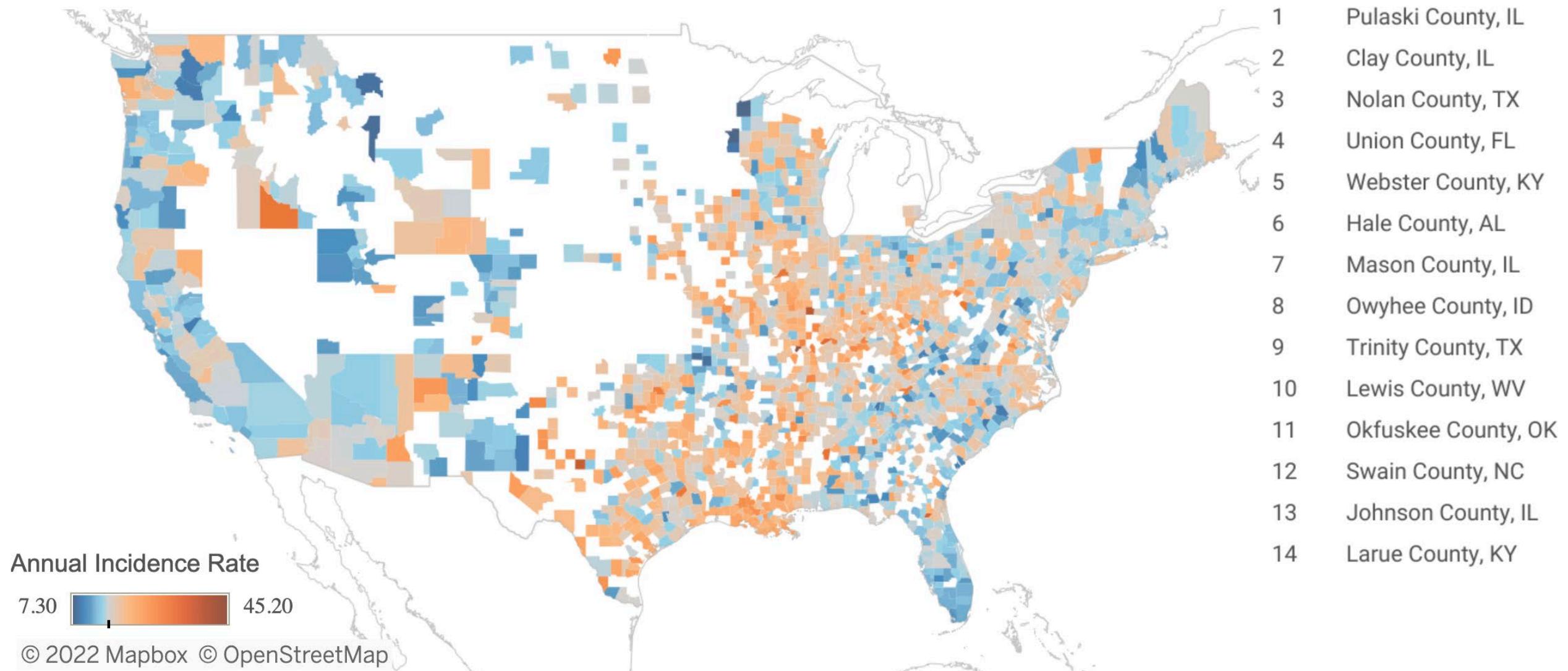
People in the west and east coast use Twitter more often than in the other regions.

Average Volume of Twitter Messages Posted in a Day

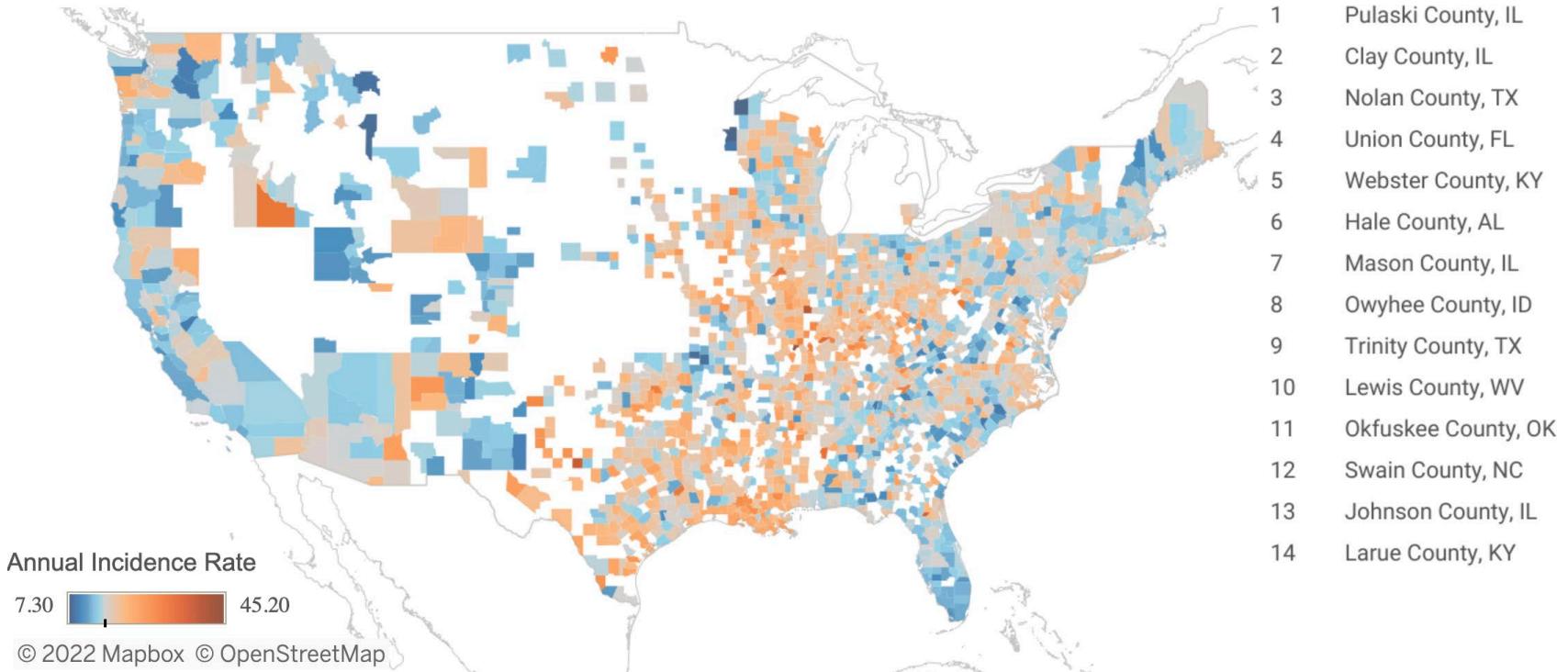


# Insensitivity to Sample Size

# Kidney Cancer and Insensitivity to Sample Size



# Kidney Cancer and Insensitivity to Sample Size

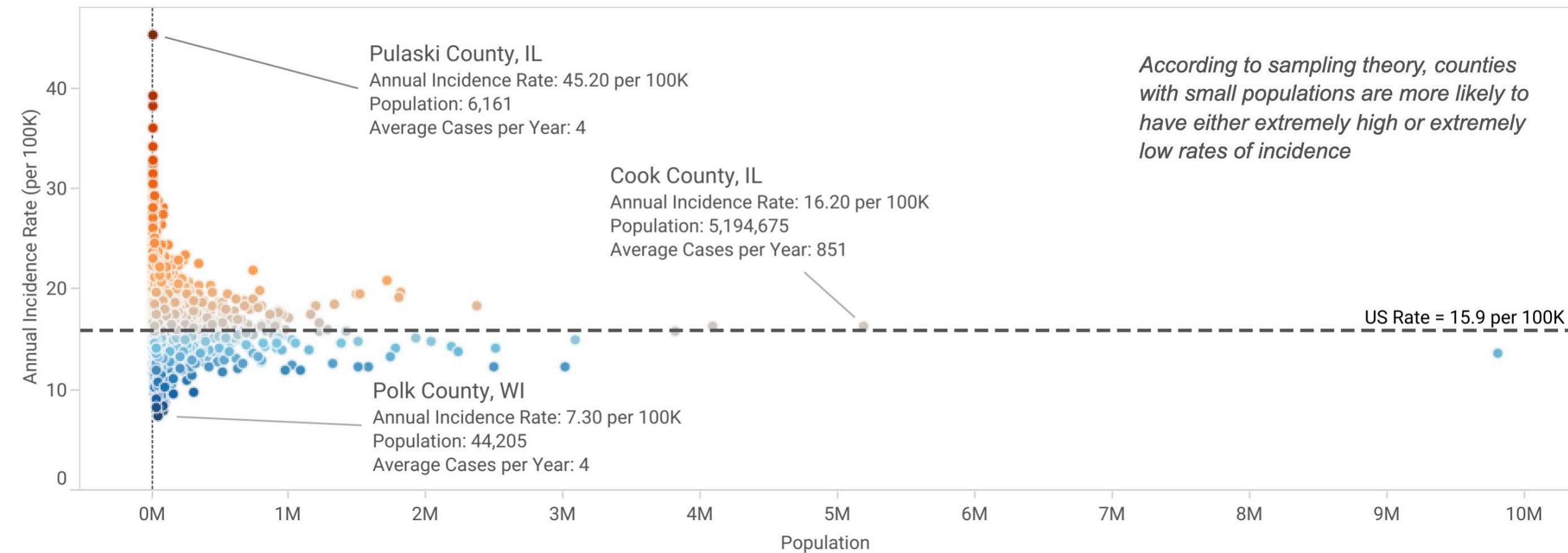


Wainer, Howard, and Harris L. Zwerling. "**Evidence that smaller schools do not improve student achievement.**" *Phi Delta Kappan* 88.4 (2006): 300-303.

*"The counties in which the incidence of kidney cancer is **lowest** are mostly rural, sparsely populated, and located in traditionally Republican states in the Midwest, the South, and the West."* p.109

*"Now consider the counties in which the incidence of kidney cancer is **highest**. These failing counties tend to be mostly rural, sparsely populated, and located in traditionally Republican states in the Midwest, the South, and the West."*

**Insensitivity to sample size** is a cognitive bias that occurs when people judge the probability of obtaining a sample statistic without respect to the sample size.

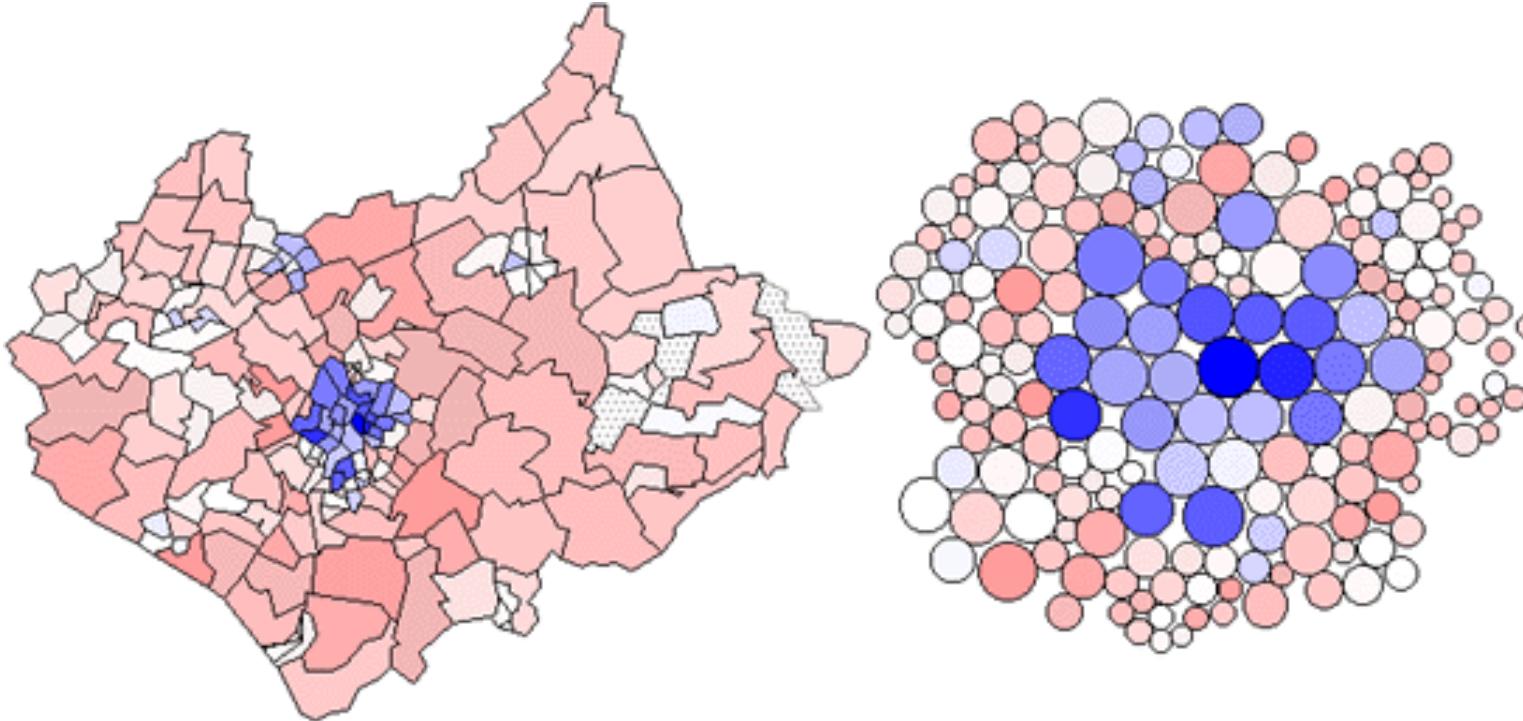


# Skewed Spatial Distributions

Small and large regions



# Dorling Cartogram



Regions are represented by symbols instead of region borders.

Symbol size is not proportional to region size.

Topological relationship is preserved approximately.

**306**

Joseph R. Biden Jr. ✓

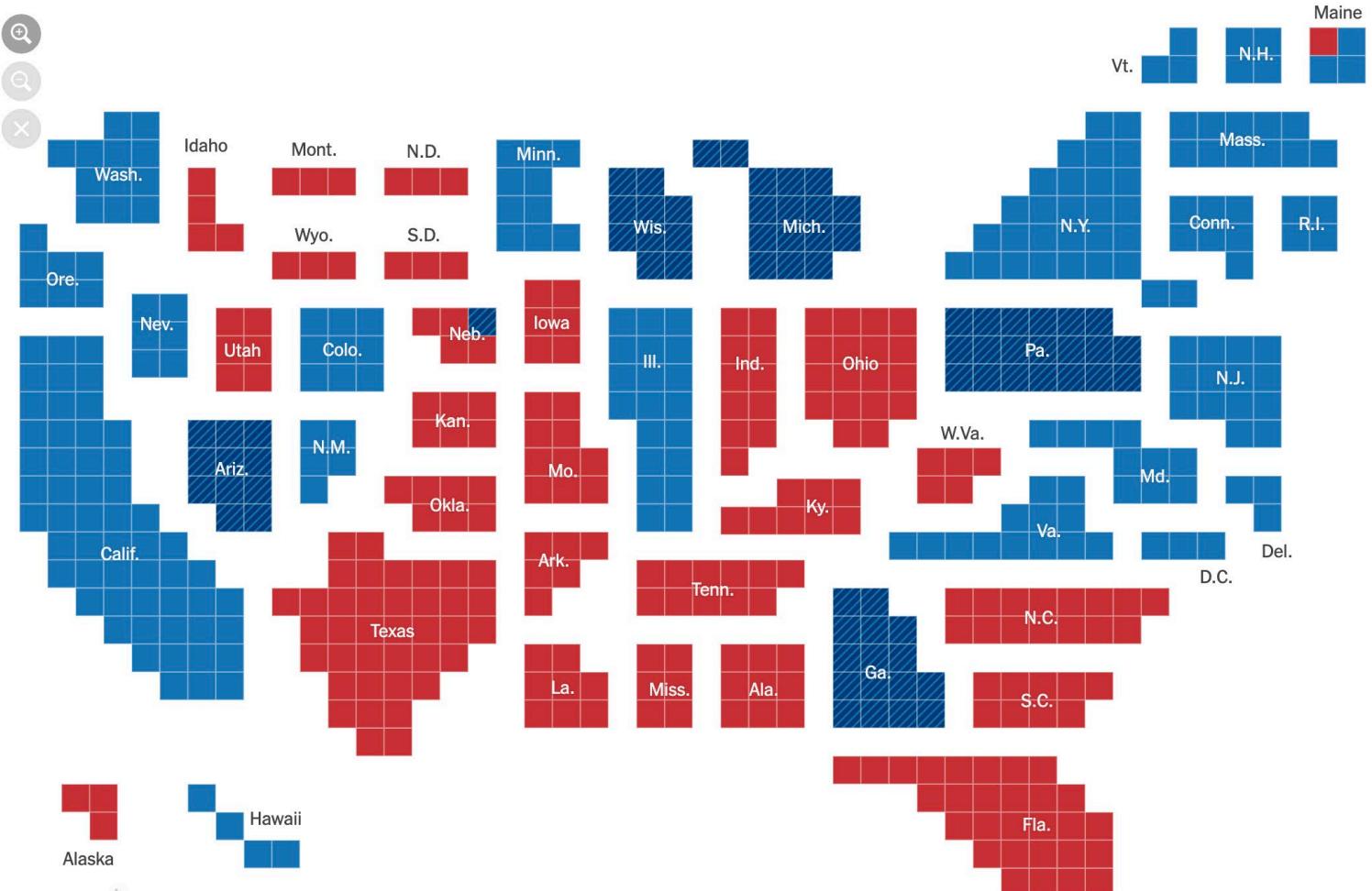
81,284,666 votes (51.3%)

**232**

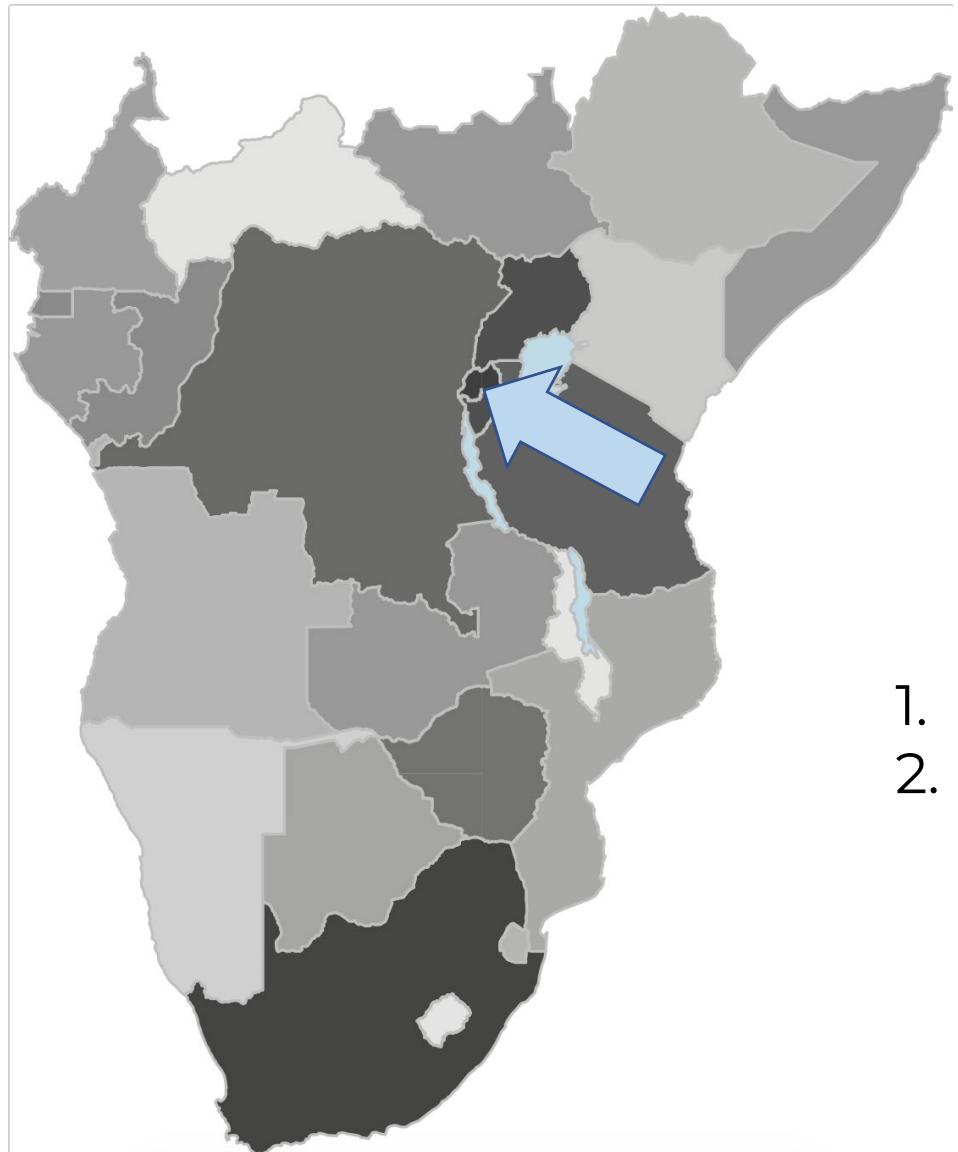
Donald J. Trump

270  
TO WIN

74,224,319 votes (46.8%)

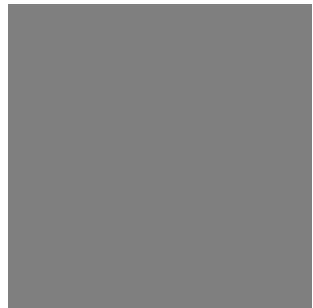
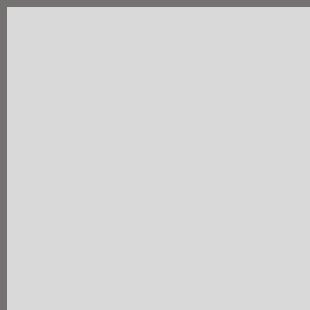
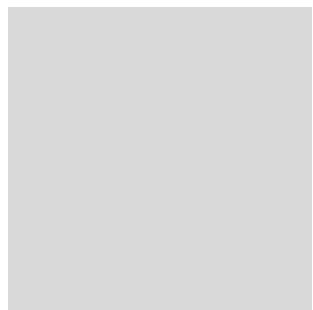


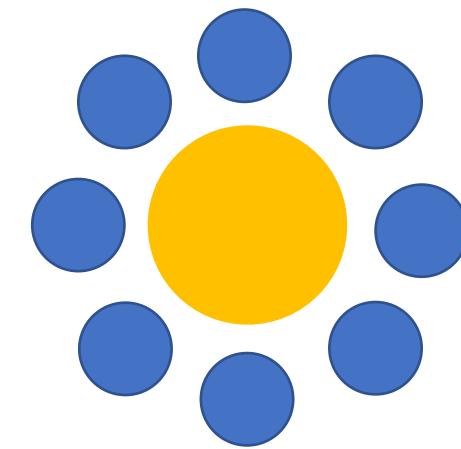
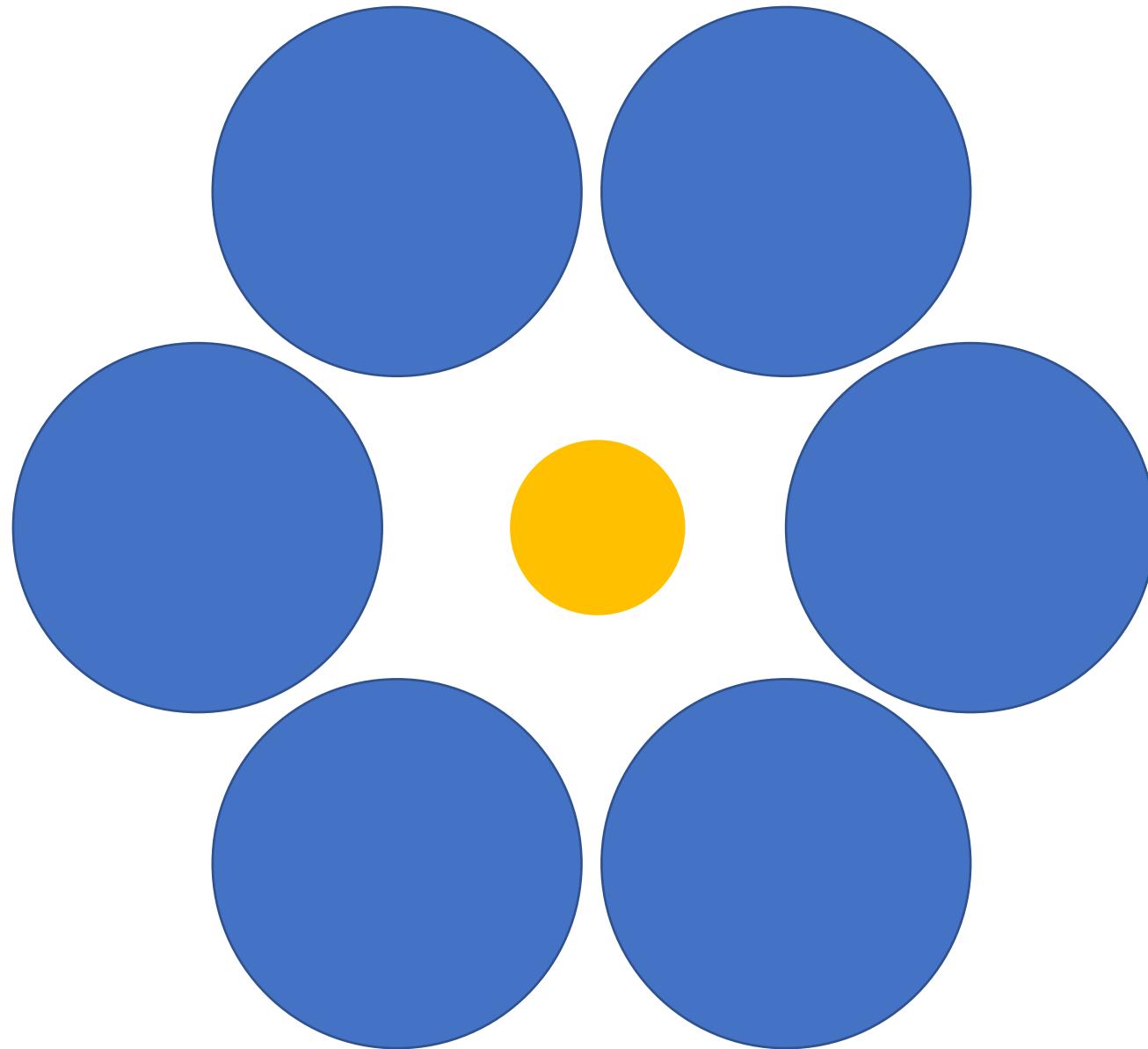
# Perception Issues



Which country in Africa has the highest value?

1. Color perception is highly affected by size.
2. Color of the surrounding regions affects the way we perceive color.

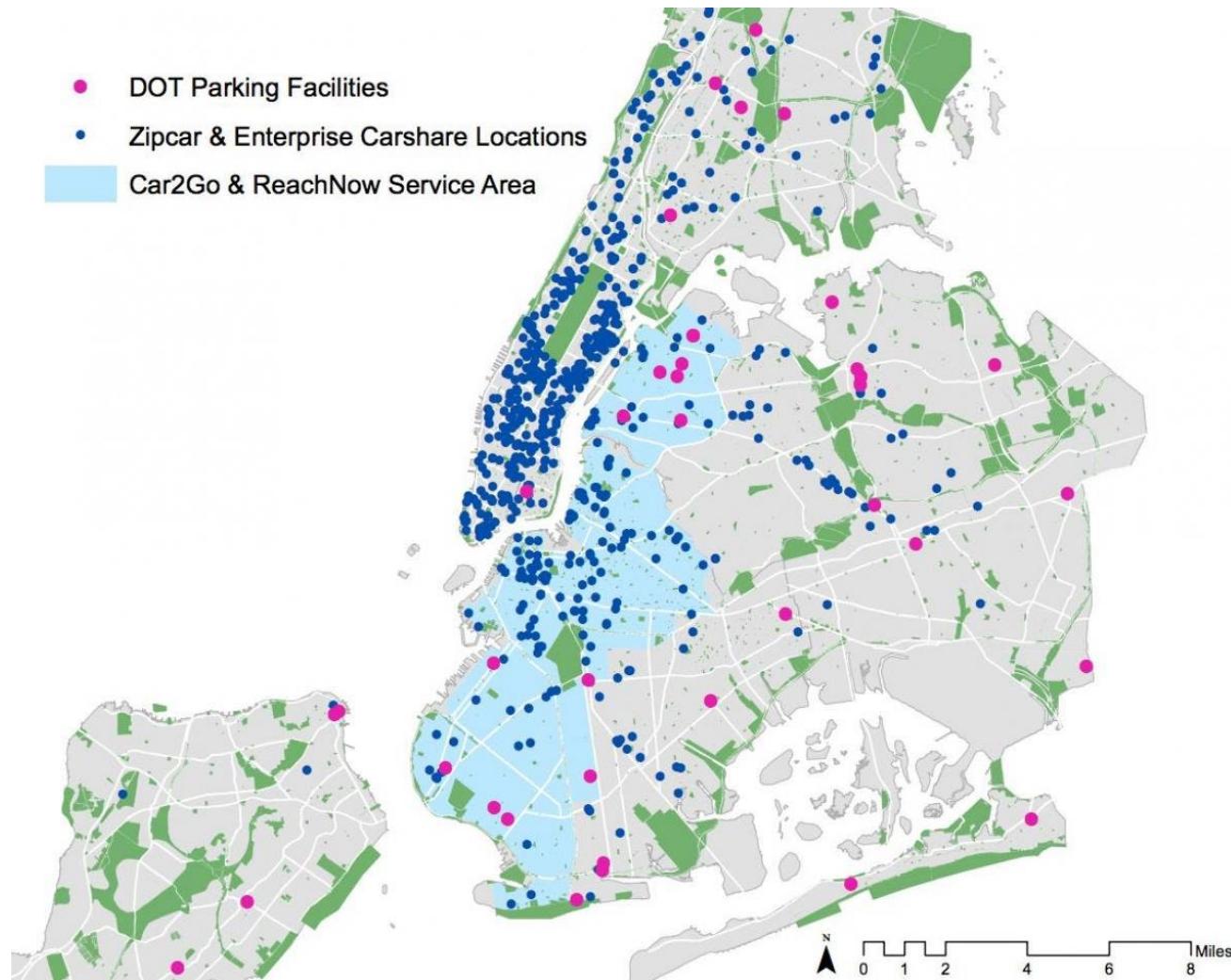




# Other Topics

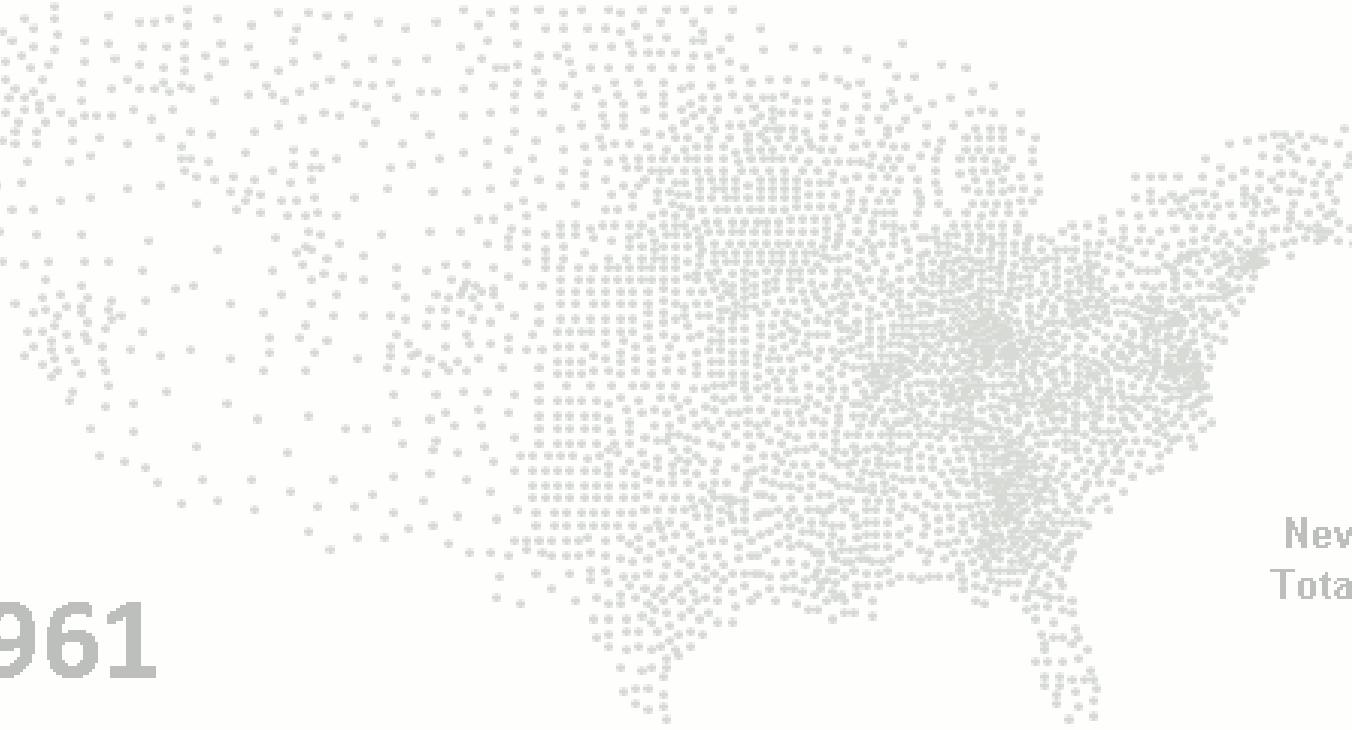
# Interference

- Use few and low saturated color for map.
- No strong lines for border.
- Include only necessary spatial features.



# Animation

1961

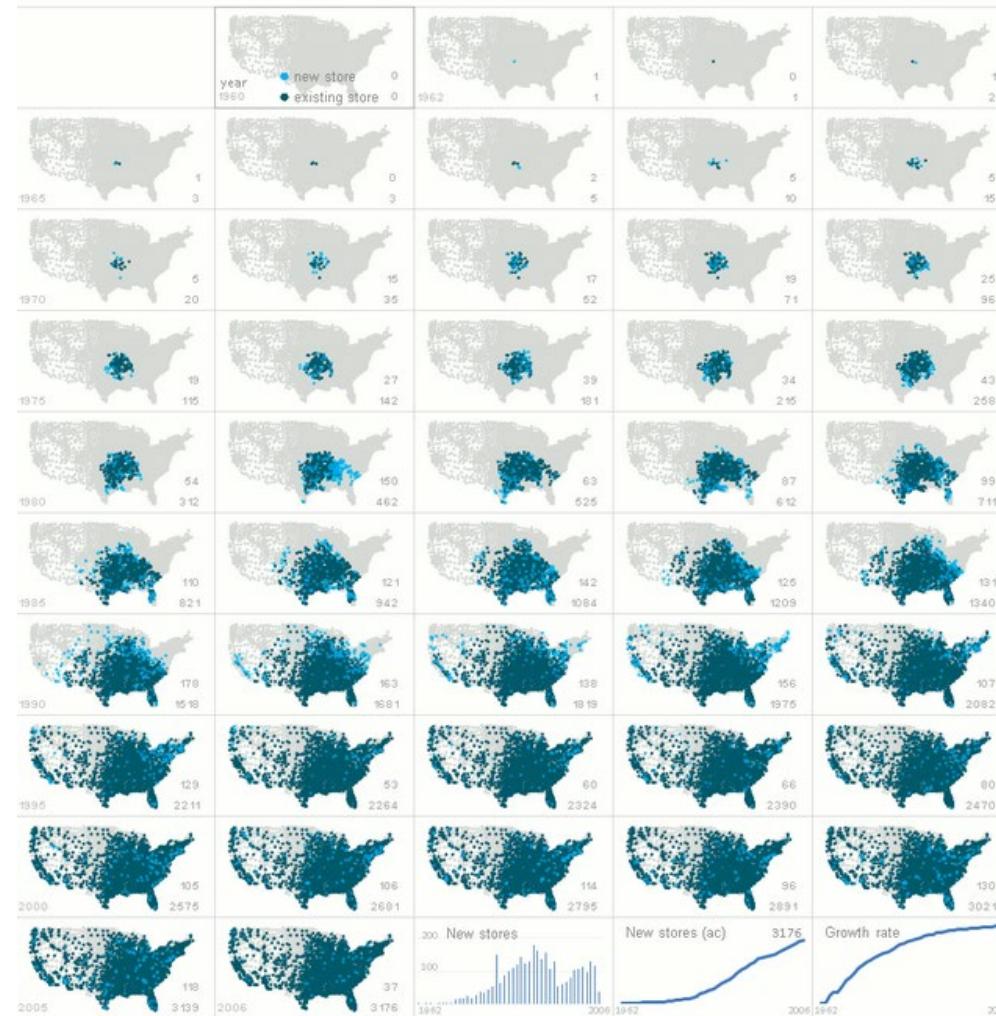


97

Source: <http://www.excelcharts.com/blog/animation-small-multiples-growth-walmart-excel-edition>

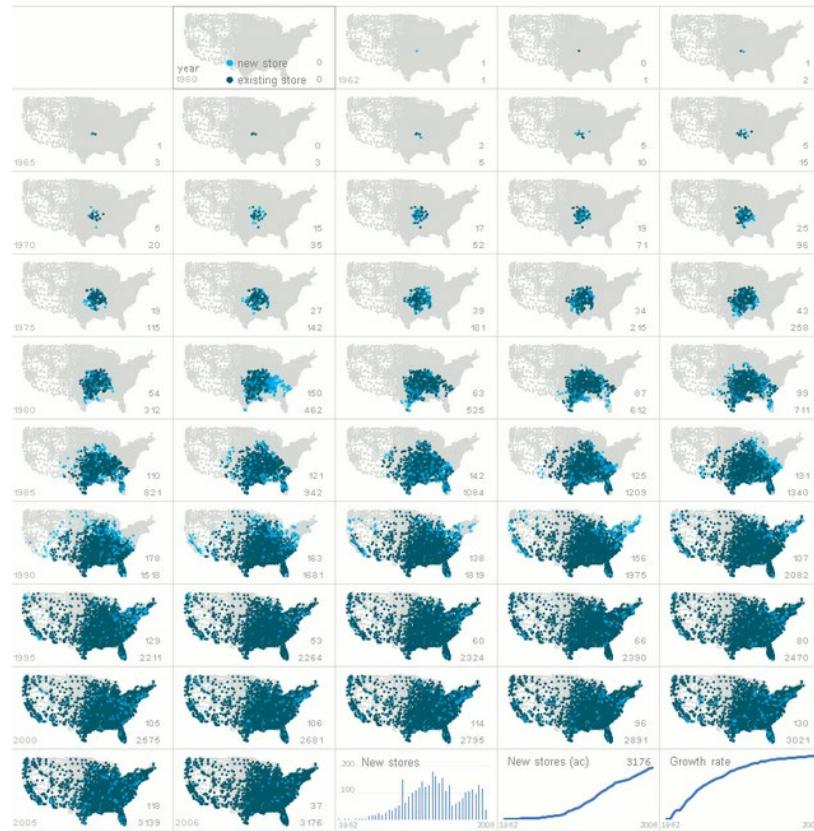
# Facet

Growth of Walmart



# Remove the map

Growth of Walmart



The Growth of Walmart: Store Openings by Year and State

