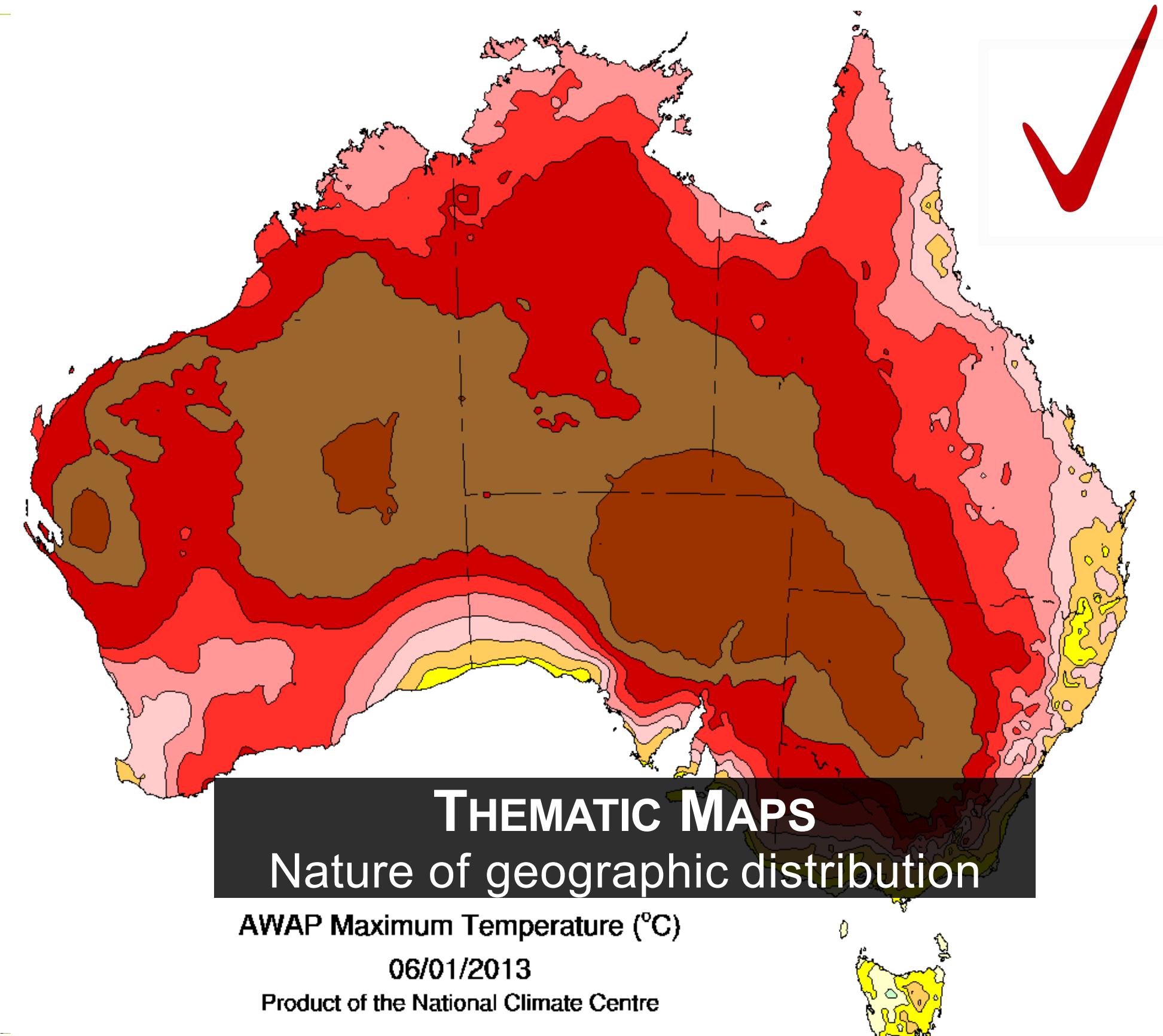


# Geographical plotting

Prof. Rossano Schifanella



# two (overlapping) categories



# Thematic maps

- **Visualize spatial distributions of data, e.g., population density**
- **Thematic maps serve three primary purposes.**
  - 1. They provide specific information about particular locations.
  - 2. They provide general information about spatial patterns.
  - 3. They can be used to compare patterns on two or more maps.

# Design is driven by

- Data
  - Categorical, ordinal, interval, ratio

		Nominal	Ordinal	Interval	Ratio
Categorical	mutual exclusive, not ordered, categories e.g., five different genotypes, average no meaning	frequency distribution.	Yes	Yes	Yes
Ordinal	order matters but not the difference e.g., movie ratings	median and percentiles.	No	Yes	Yes
Interval	difference between two values is meaningful e.g., temperatures in Celsius, a temperature of 100 degrees C is not twice as hot as 50 degrees C	add or subtract.	No	No	Yes
Ratio	as interval but has a clear definition of 0.0 e.g., temperature in Kelvin,	mean, standard deviation, standard error of the mean.	No	No	Yes
	ratio, or coefficient of variation.		No	No	Yes

# Design is driven by

- **Data**
  - Categorical, ordinal, interval, ratio
- **Spatial scale and granularity**
  - discrete vs continuous

## Discrete

only found at fixed locations or when the data represent only specific values, e.g., # accidents at crossings

Point Line Polygon

## Continuous

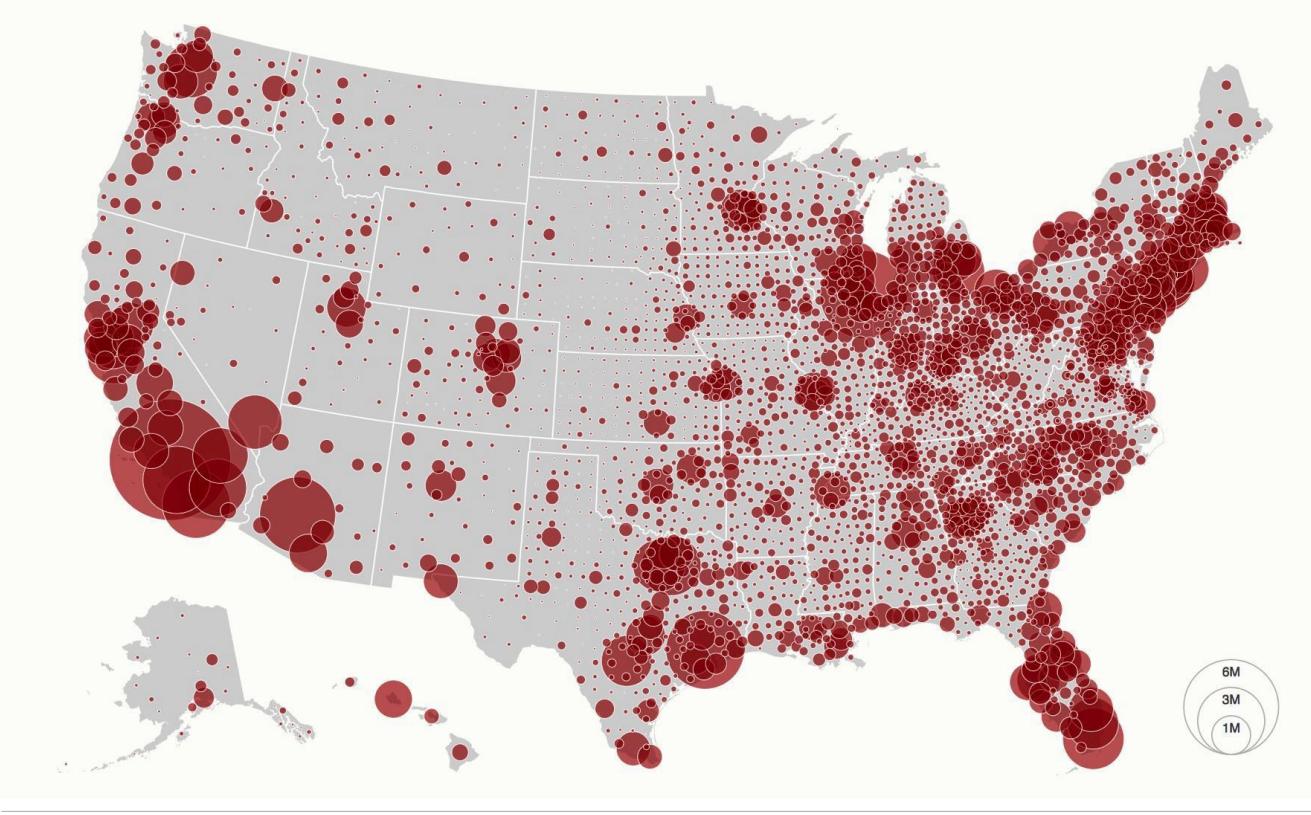
seen throughout the mapped area and smoothly transitions from one value to another, e.g., air temperature.

Surface Volume

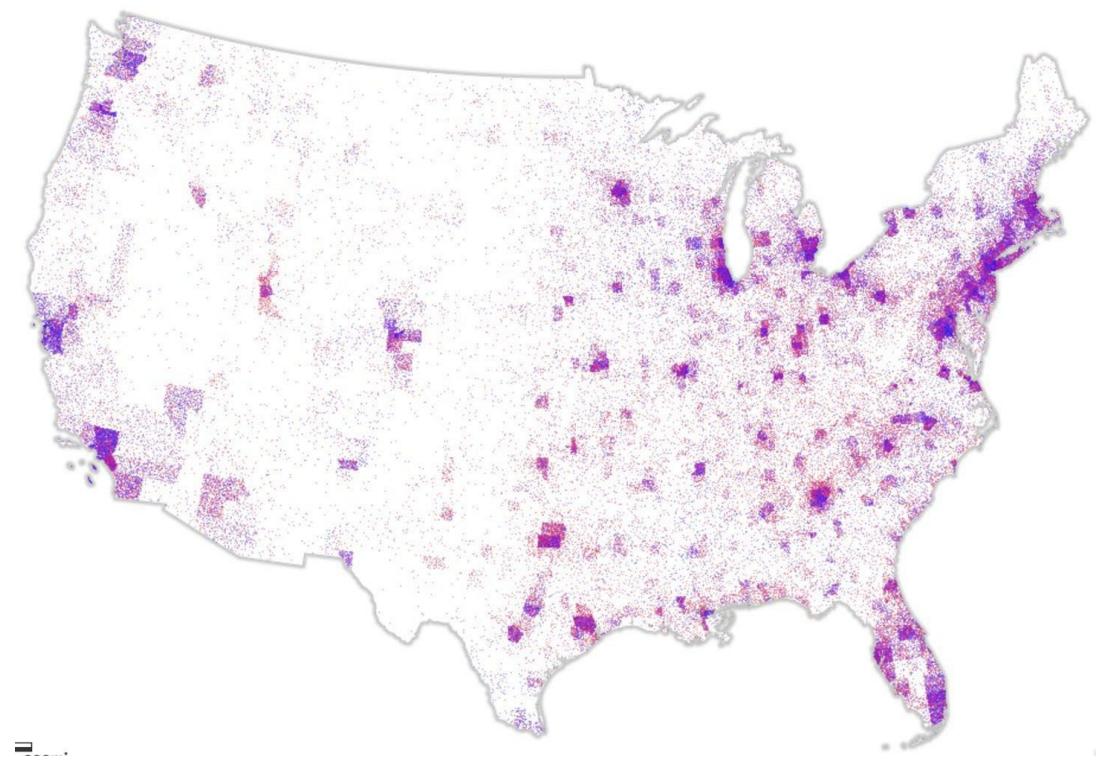
# Design is driven by

- **Data**
  - Categorical, ordinal, interval, ratio
- **Spatial scale and granularity**
  - discrete vs continuous
- **Human visual perception and aesthetics**
  - choosing the correct visual variables, e.g., symbols, colors
- **Audience**
  - knowing who will read the thematic map and for what purpose helps define how it should be designed
    - political scientist vs biologist

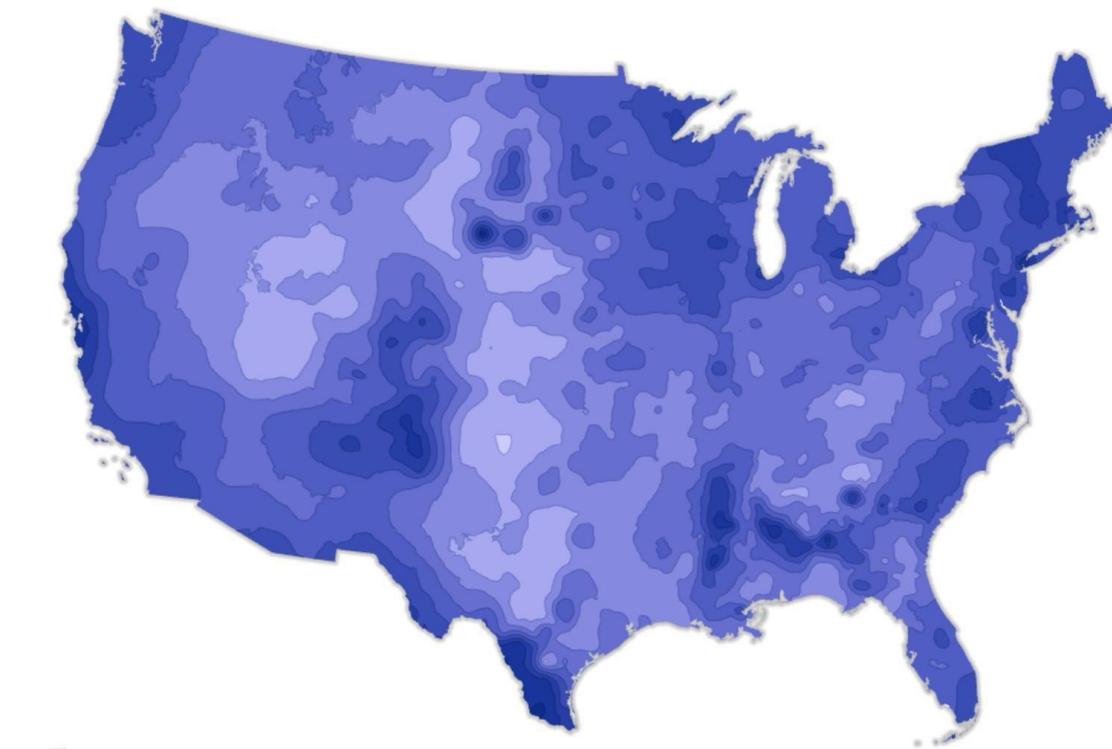
## Proportional symbols



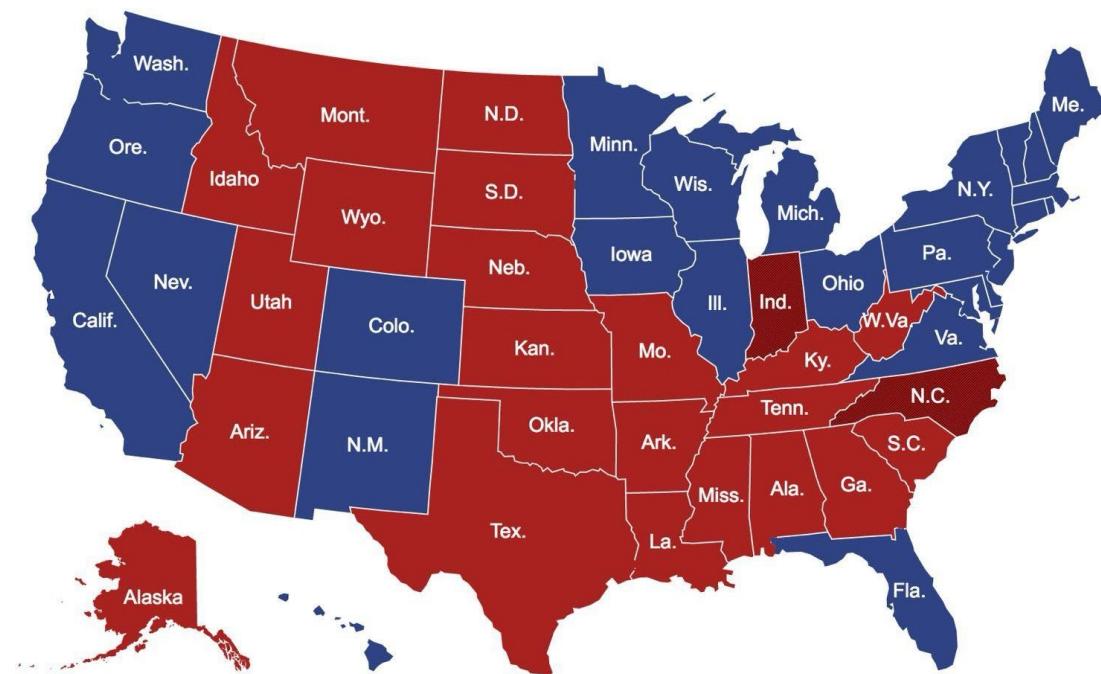
## Dot distribution



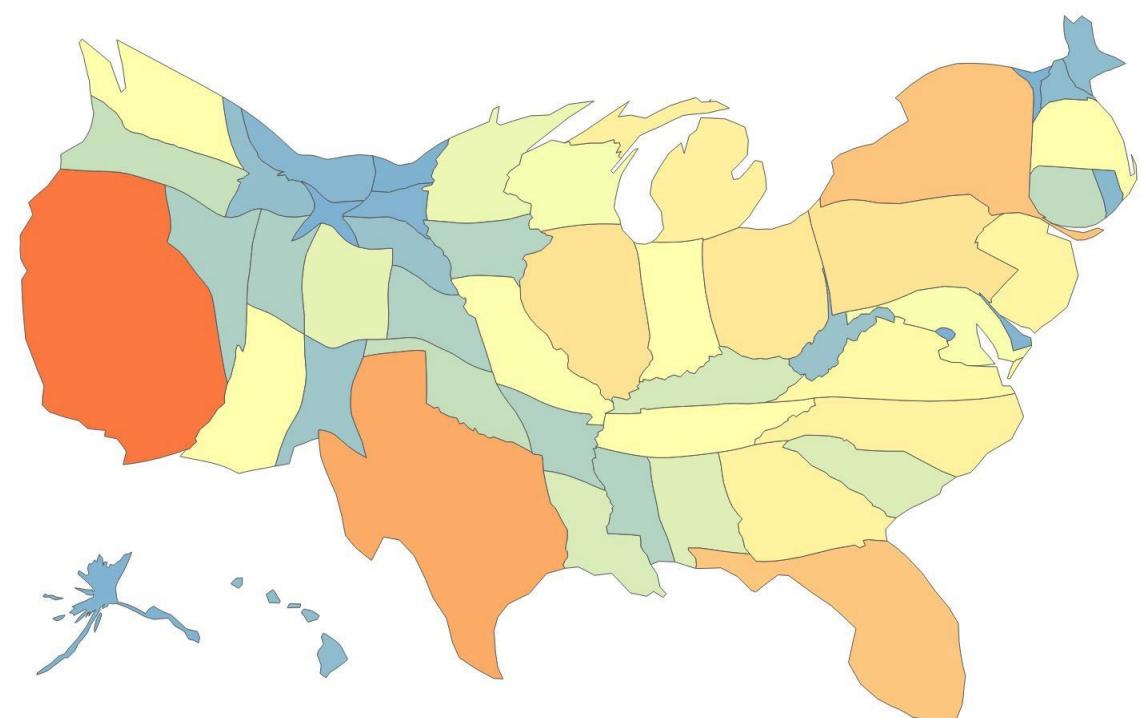
## Isopleth



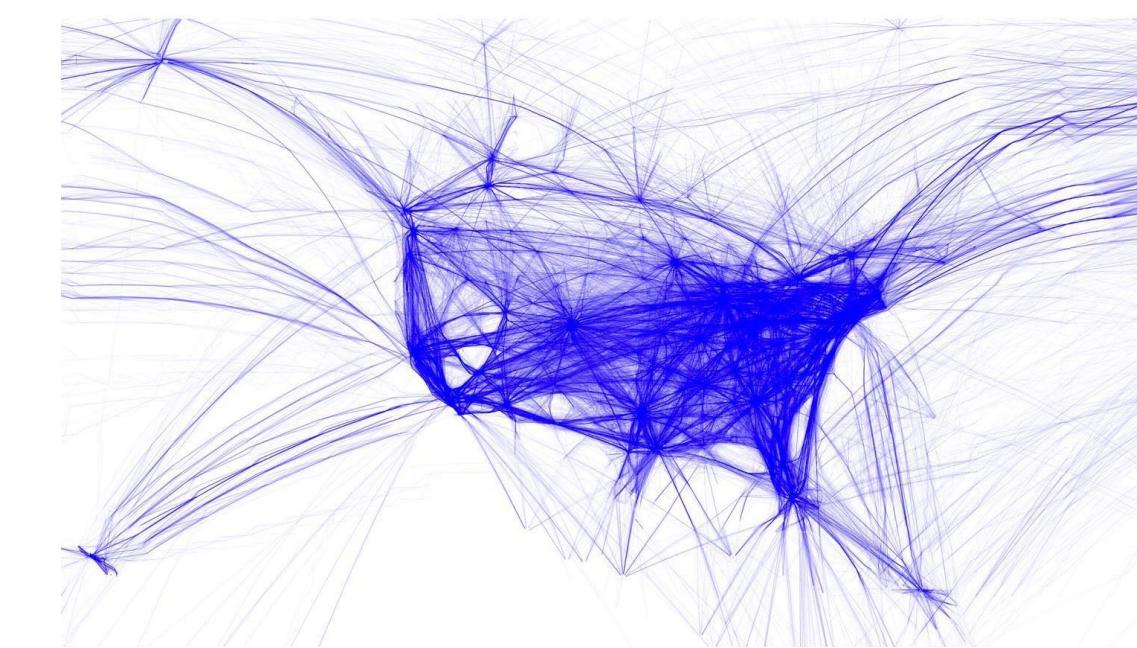
## Choropleth



## Cartograms

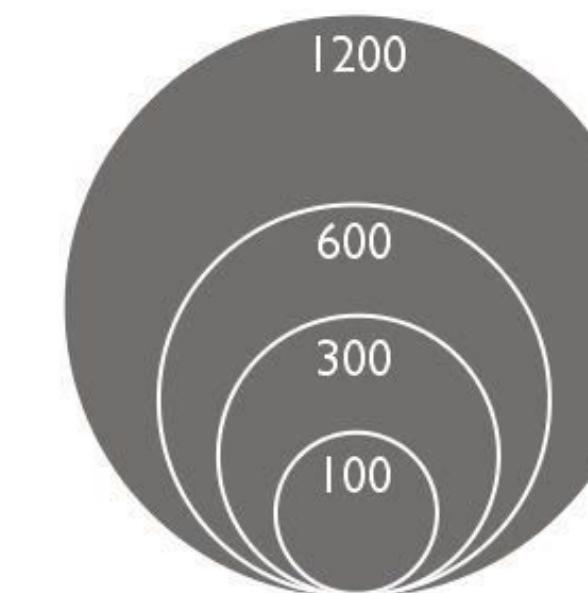


## Flow Maps

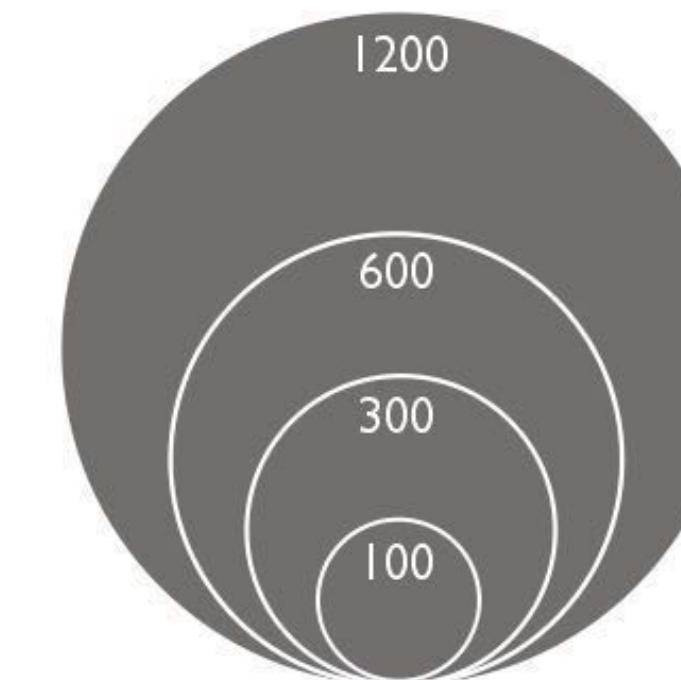


# Proportional symbol maps

- Represent data variables by symbols that are sized, colored according to their amount or type.
- Data is (or can be) aggregated at points within areas.
- Different methods for setting symbol size:
  - absolute scaling
  - apparent magnitude (perceptual) scaling
- psychophysical research revealed that people tend to correctly estimate lengths, and to underestimate areas and volumes.
  - range grading



Absolute Scaling



Apparent Scaling  
(Flannery's Compensation)

## Proportional Symbol

2012 US Presidential election results by County, by total votes

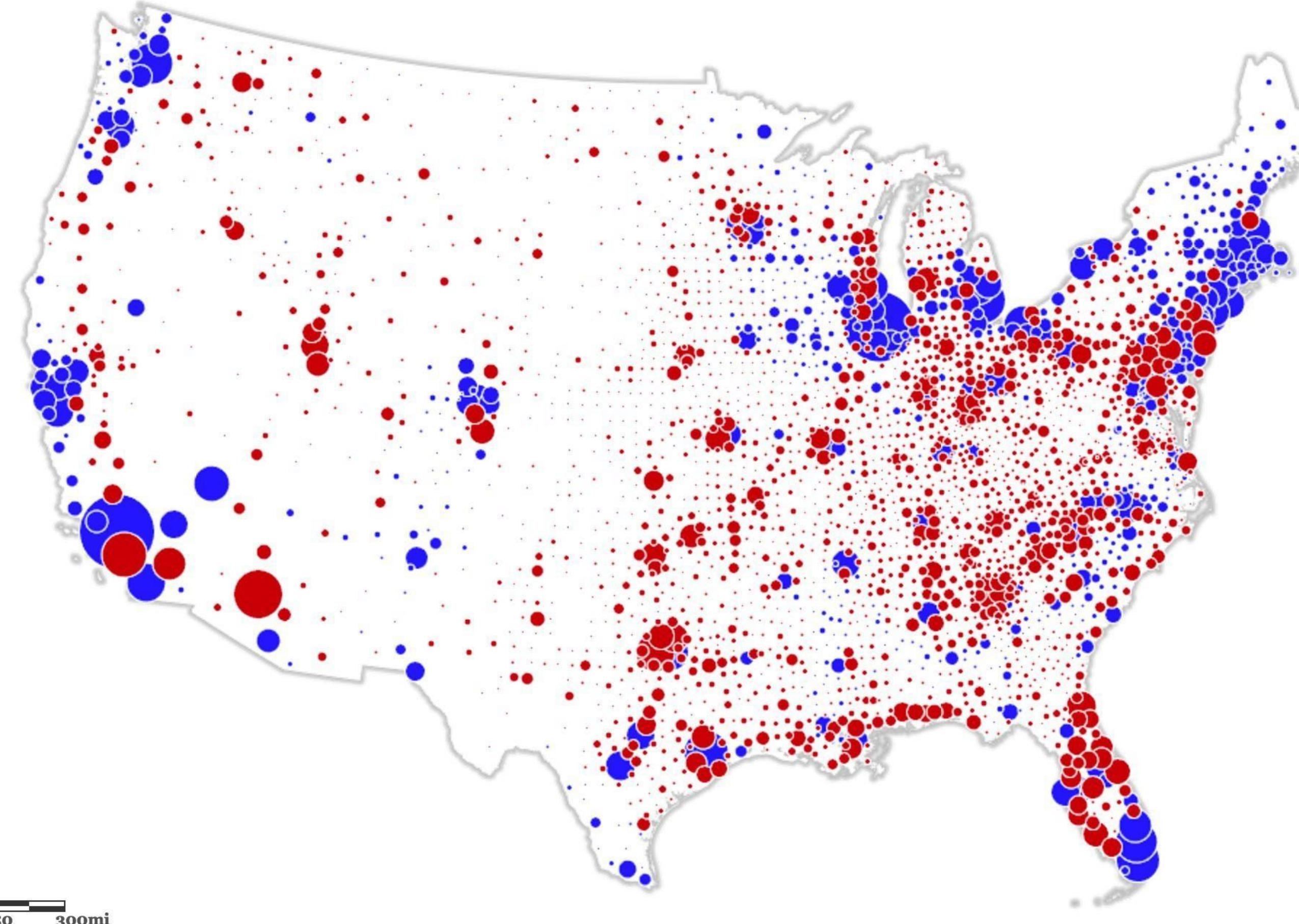
### Map type

The purpose of a **proportional symbol** thematic map is to show how features differ in quantity for the theme being mapped. In this example of the 2012 Presidential election, the map is designed to show the number of votes cast for the predominant party in each County.



### Data

For the theme being mapped, the data should be **numerical (quantitative)** and represent differences between features on an **interval or ratio** scale of measurement. The map type requires data to be absolute, as totals. Here, the vote totals are augmented by symbols that define a second **categorical** characteristic of the data, namely 'Republican' or 'Democrat'.



### Symbols

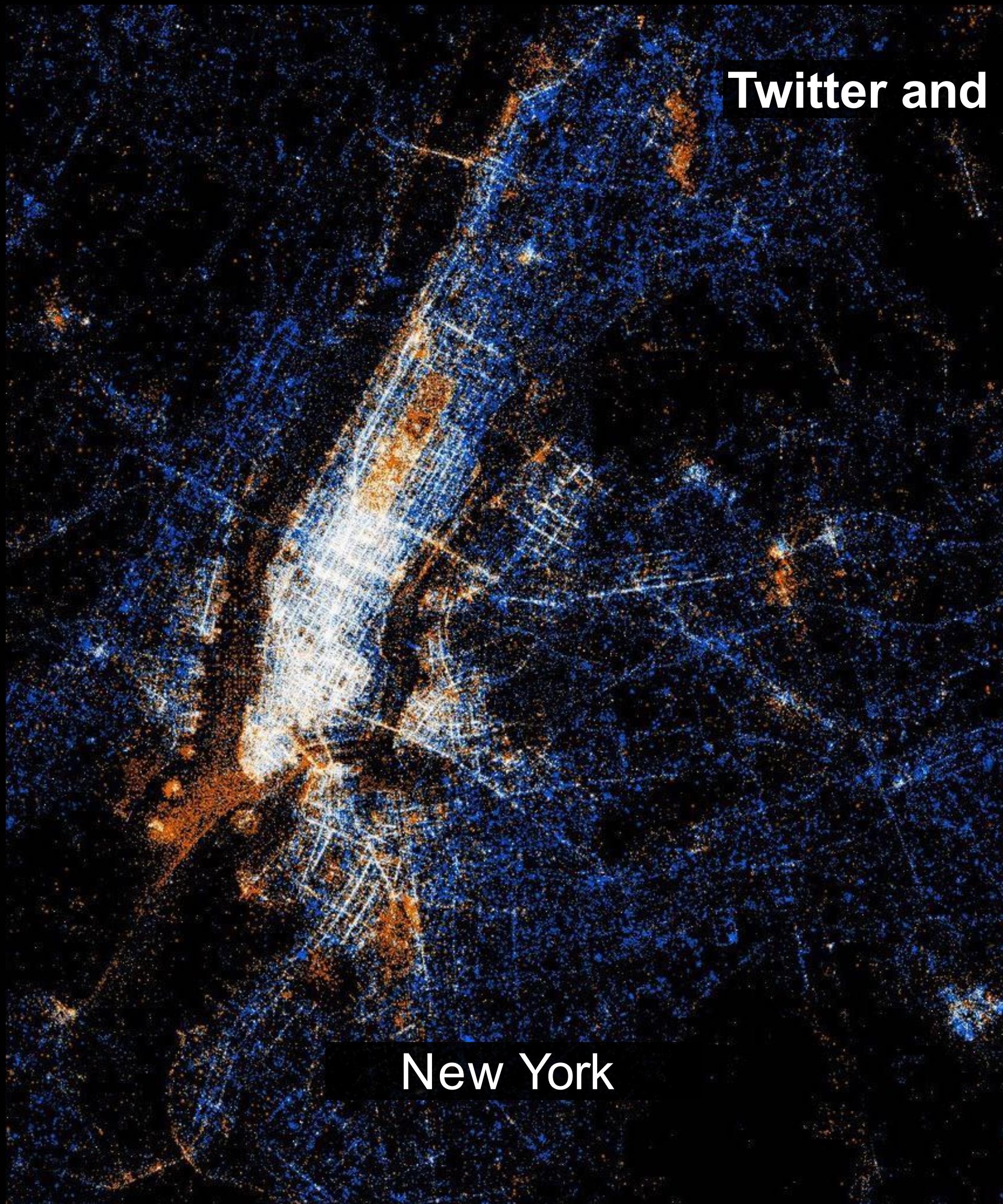
Symbols are scaled to the data values and should be designed so that different magnitudes of data can be easily distinguished from one another through variation in the **size** of the symbol, used as an **ordering visual variable**. Symbols should be scaled so that the smallest are visible and the largest do not overly smother the map.

Kenneth Field, politico.com

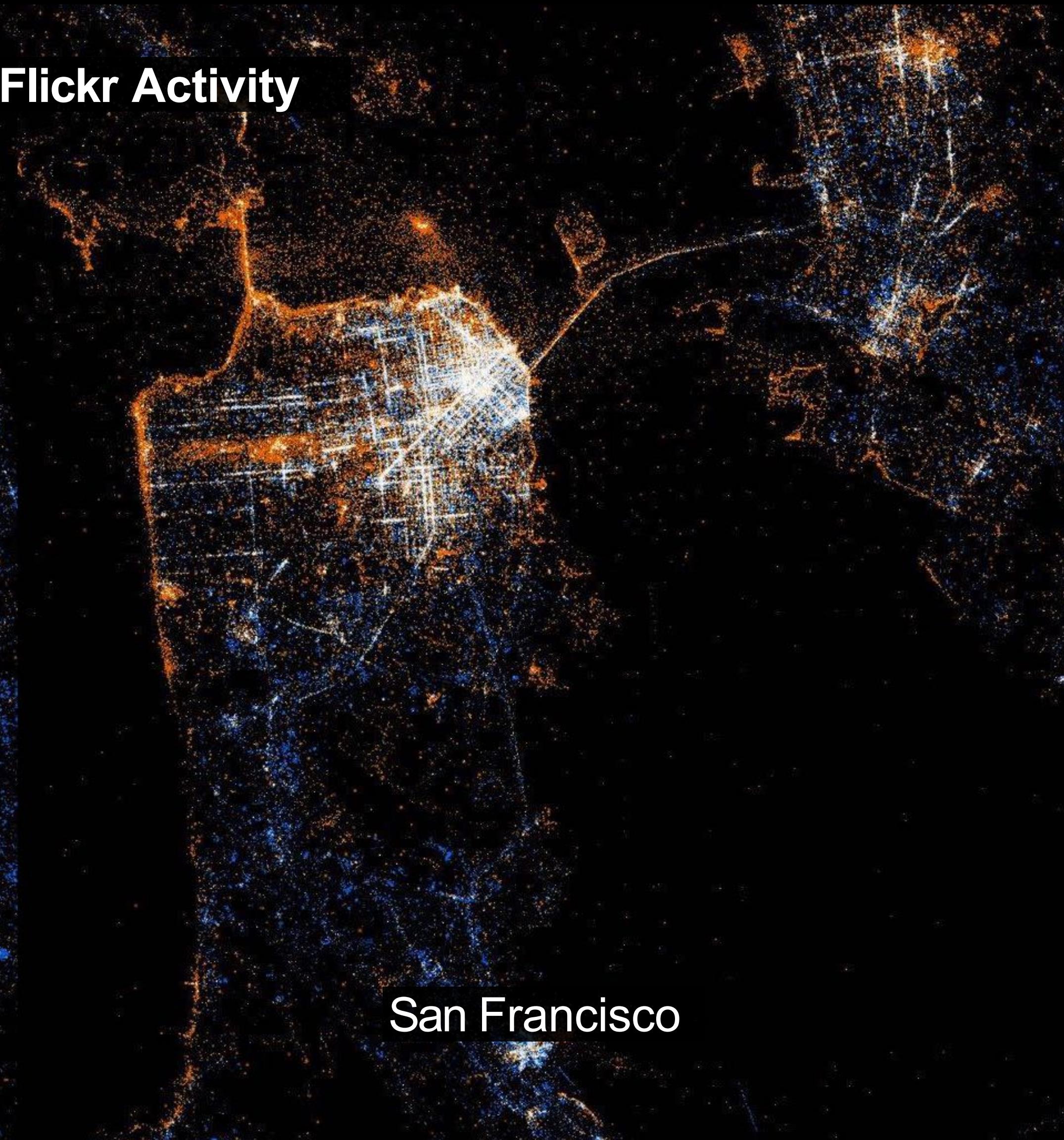


# dot distribution

Twitter and Flickr Activity

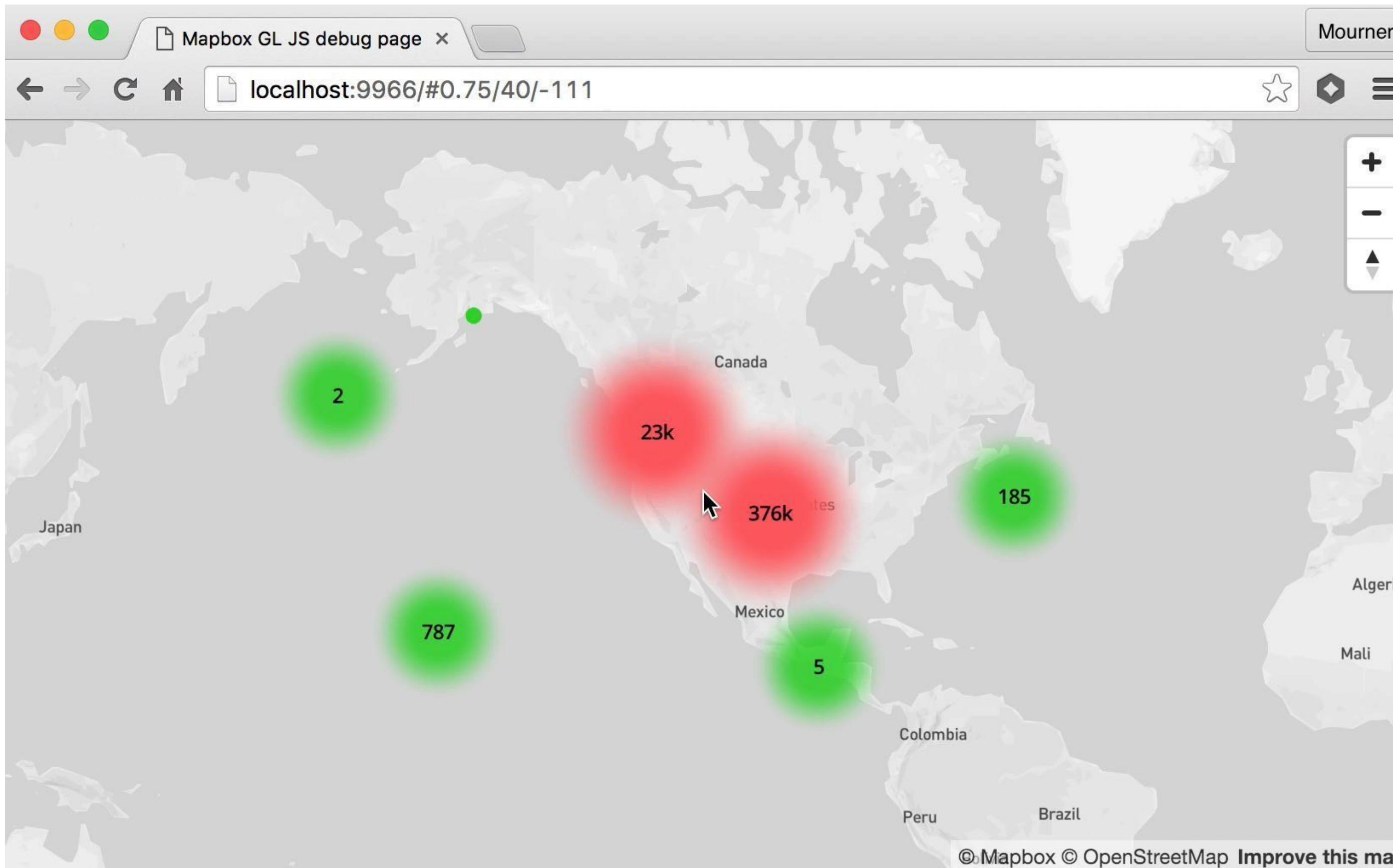


New York

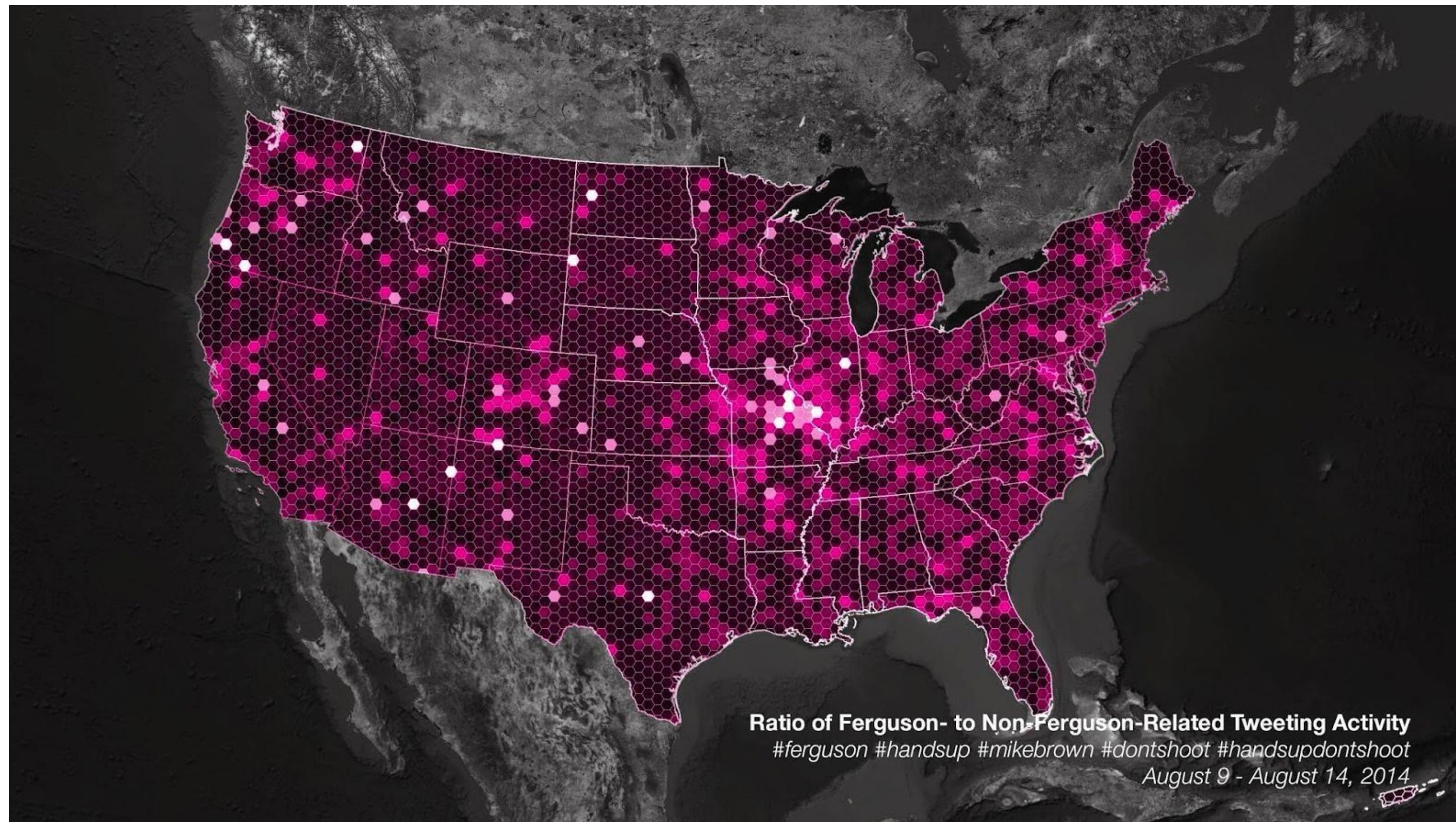


San Francisco

# Clustering (dealing with a lot of points)

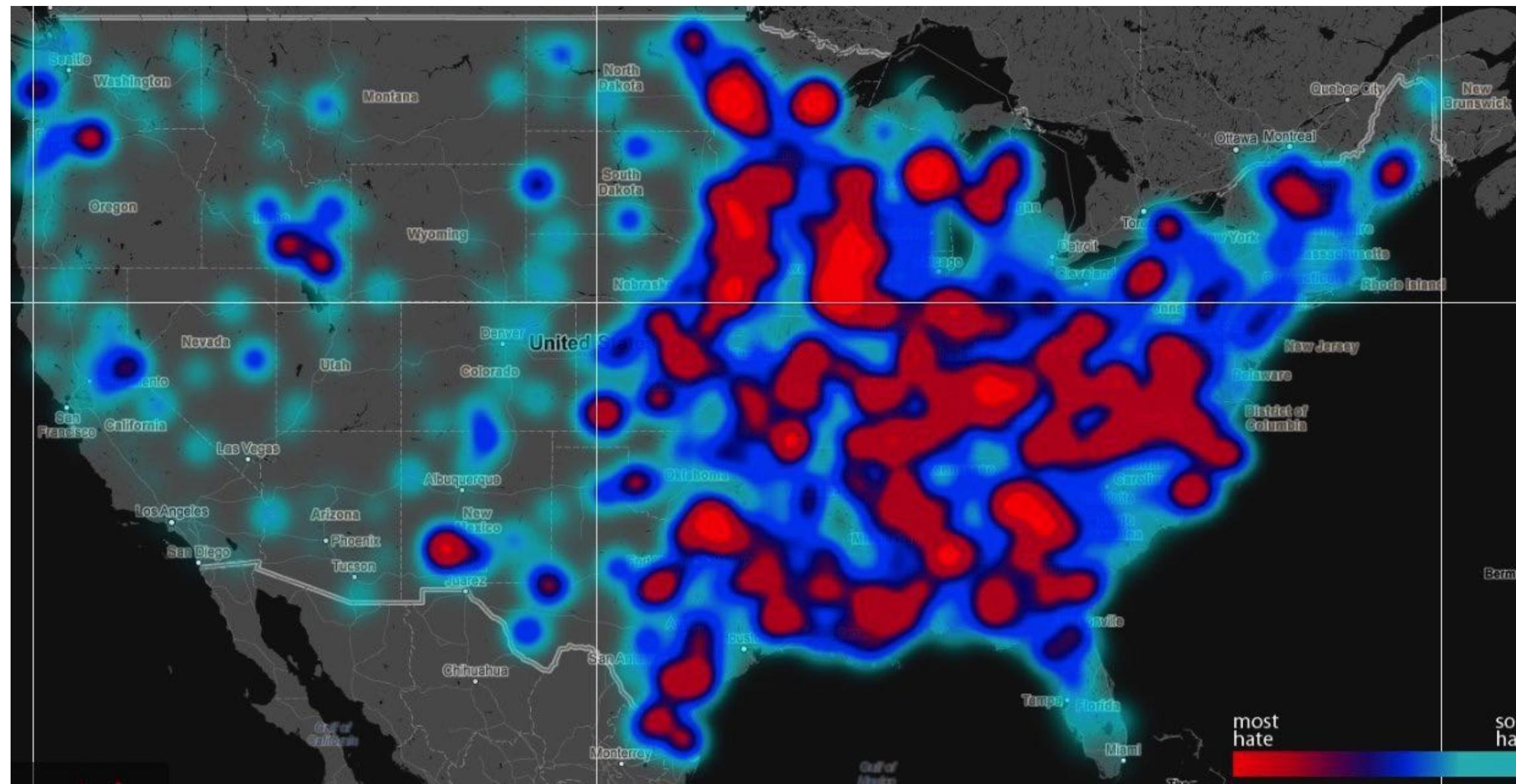


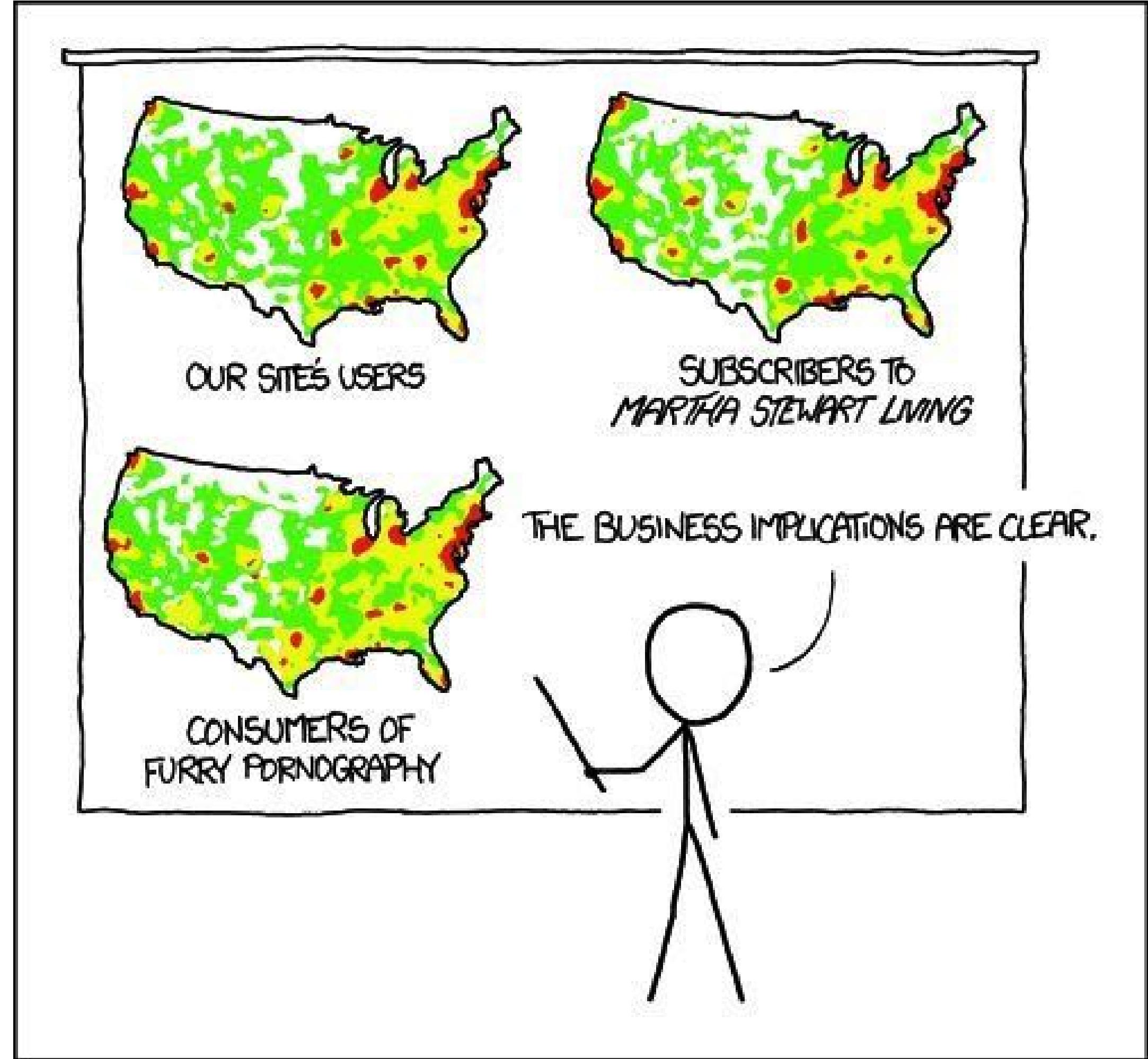
# Hexbins



# Heatmaps

- Used to identify clusters where there is a high concentration of activity (attribute under analysis)
- They can ben also useful for doing hotspot analysis.





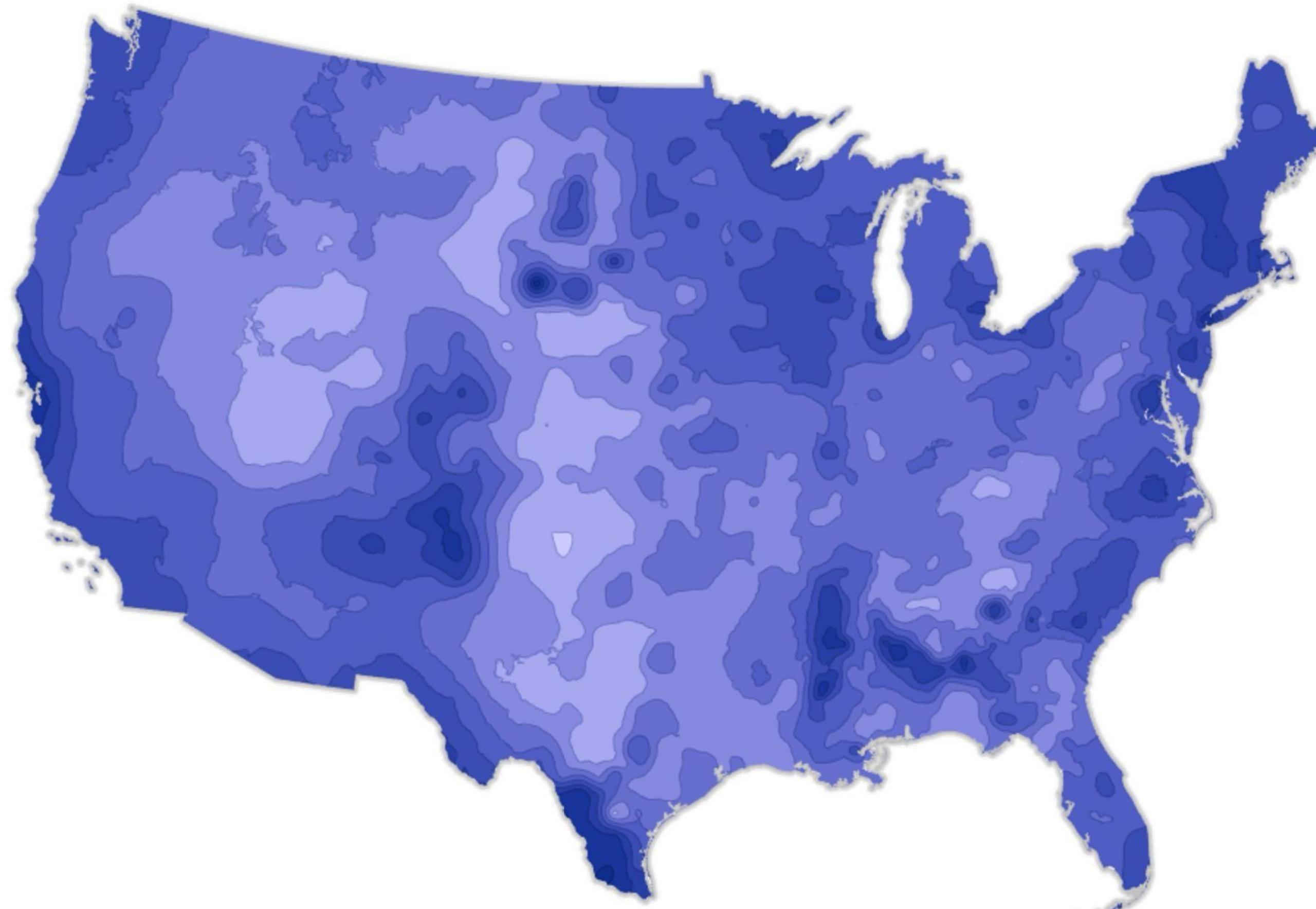
PET PEEVE #208:  
GEOGRAPHIC PROFILE MAPS WHICH ARE  
BASICALLY JUST POPULATION MAPS

## Isopleth: Filled Contours

2012 US Presidential election results: Democrat share of vote

### Map type

An **isarithmic** map is a two-dimensional representation of a three-dimensional volume. Two types exist: an **isometric** form that is constructed from data at points and an **isoplethic** form constructed from data that occur over geographic areas. The purpose of an **isopleth** thematic map is to show how features differ in quantity as a surface. This can be achieved through representing the volume using **contour lines** or by using **filled contours** that are shaded according to the quantitative value being mapped. In this example of the 2012 Presidential election, the map is designed to show the share of the vote gained by the Democrat party based on County level data.



### Data

**Isopleth** maps are generated from data that occur over geographical areas and values represent **numerical (quantitative)** differences between features on an **interval or ratio** scale of measurement. Absolute values cannot be illustrated isoplethically due to the inherent problems of using totals for areas that might vary in size or which contain an unequal denominator of the data being mapped. This is the issue that prevents **choropleths** from being used to map totals and the same occurs for

Kenneth Field, politico.com | Kenneth Field



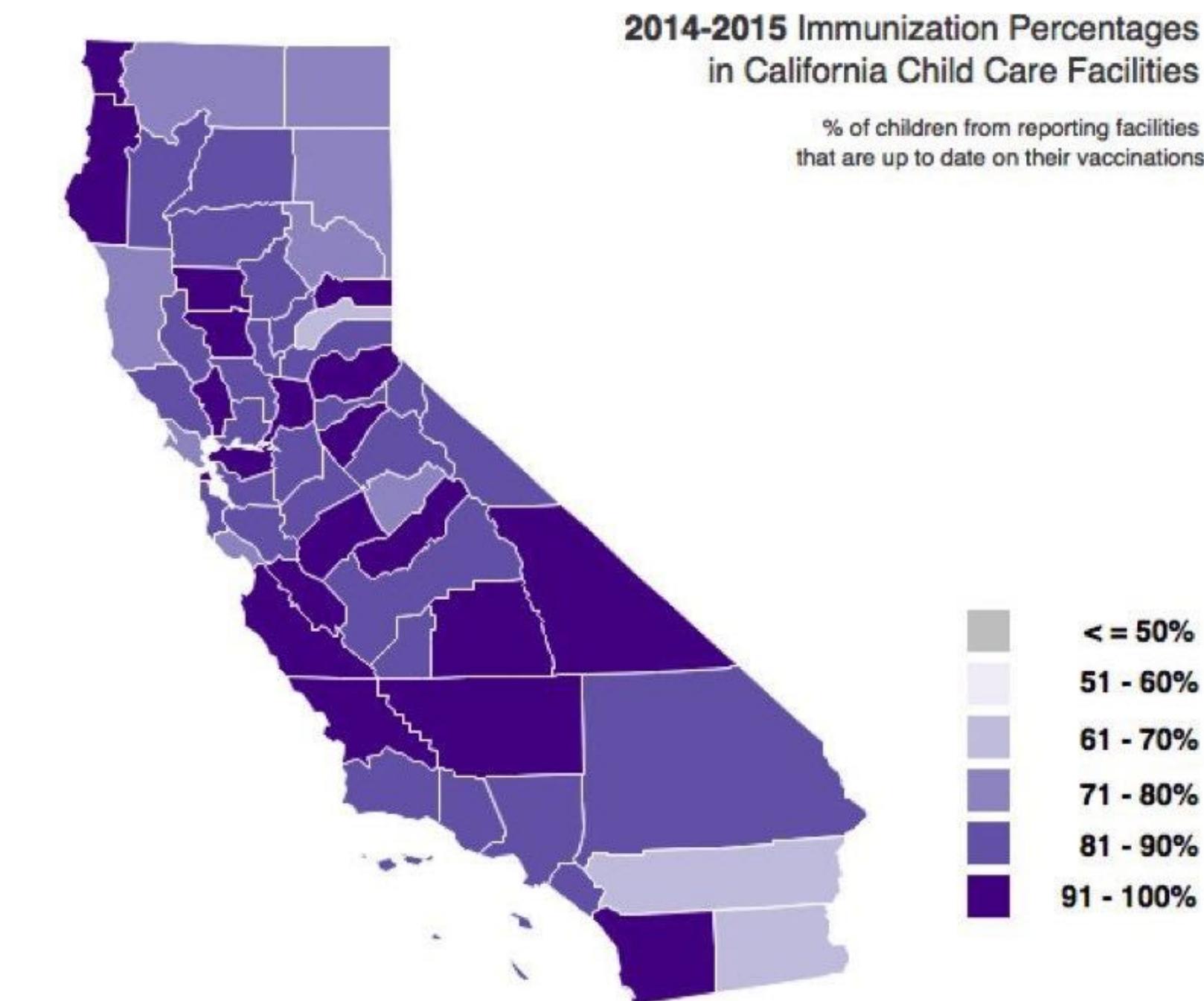
# Choropleth

(from Greek χῶρος ("area/region") + πλῆθος ("multitude"))

- Areas are shaded or colored in proportion to the measurement of the statistical variable being displayed on the map.

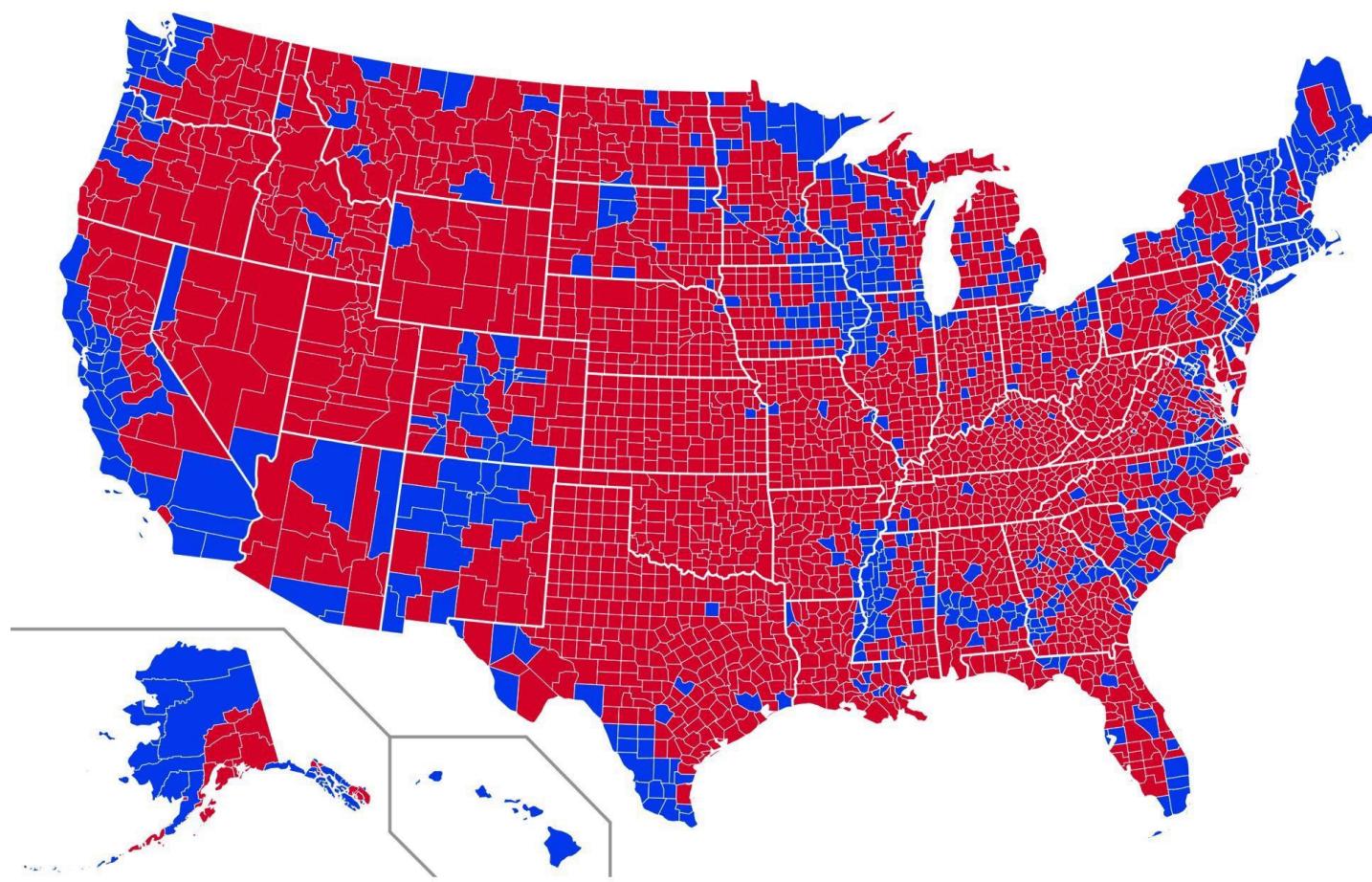
- Key factors:

- Resolution of the base map
- Data
  - source and processing
  - classification
  - MAUP
  - legend
- Symbolization



# Data type

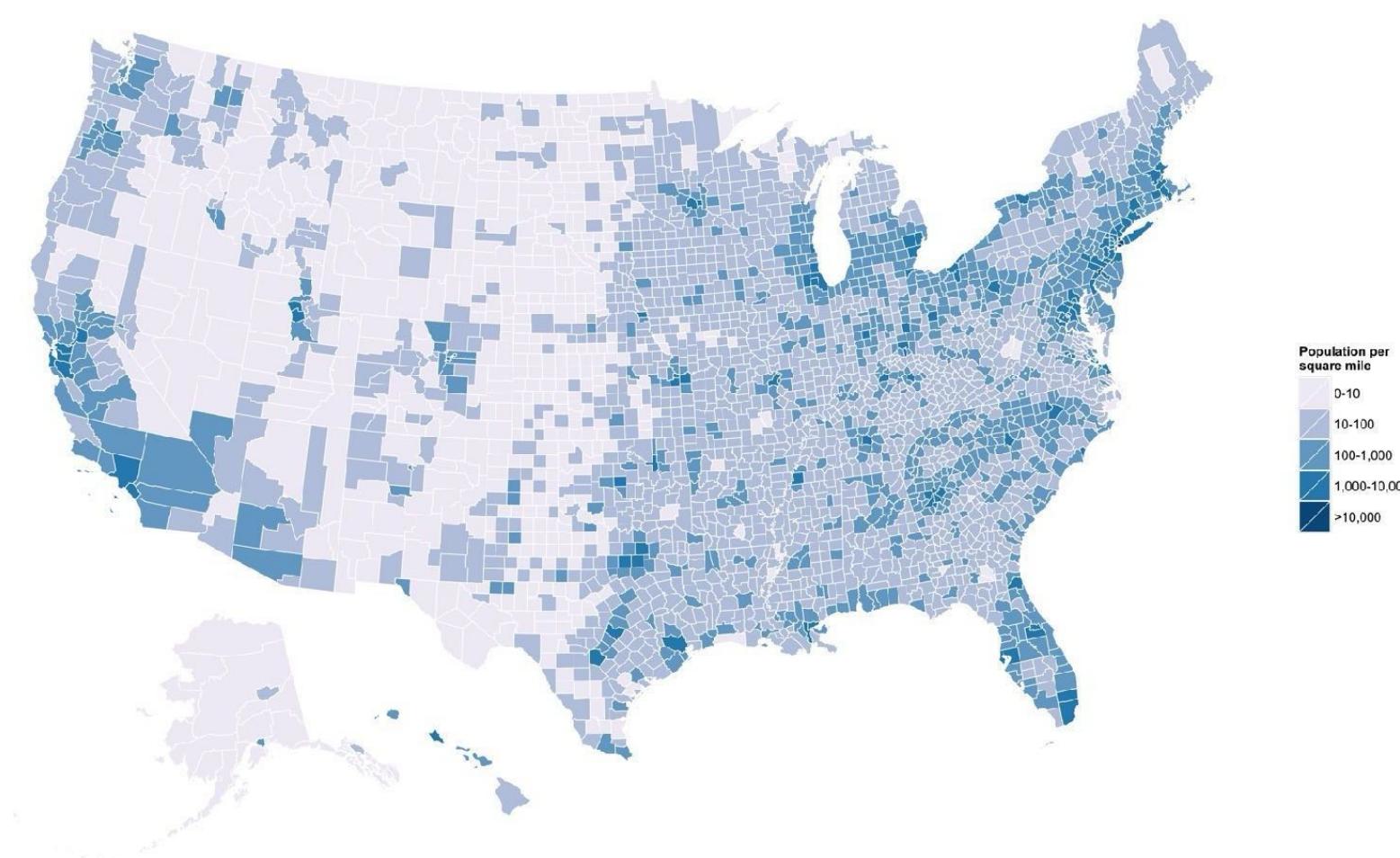
Presidential election 2008



CATEGORICAL

Obama or Romney

Population density 2014



CONTINUOUS

interval [0, 1]

Data Type

**Continuous**  
(sequential)

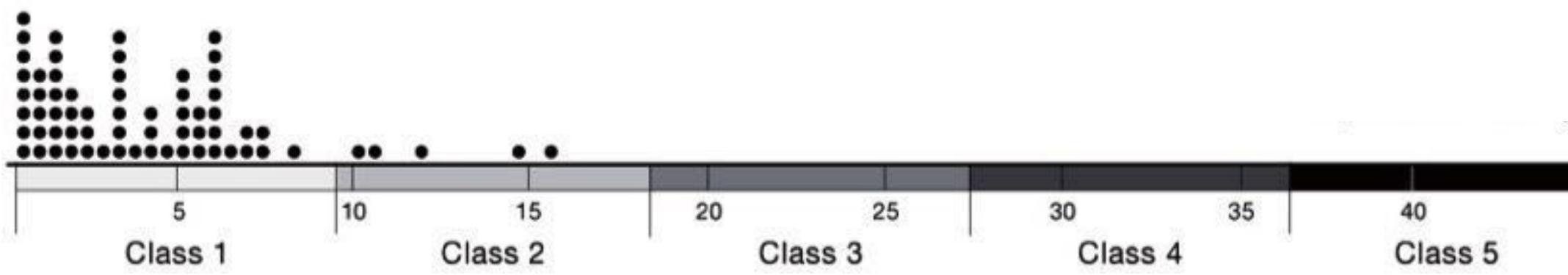
!=

Color Scheme

**Categorical**  
(qualitative)

# Classification

- Take observations and group them into data ranges or classes



How many classes?

$$5-7 \pm 2$$

George Miller (1956)  
short term memory capacity

Which method?

# Classification methods

- **Natural breaks**
- **Equal intervals**
  - not valid if your data is skewed or in presence of outliers.
- **Quantiles**
  - can position elements in a class even if being closer to the adjacent
- **Standard deviation**
- **Fisher-Jenks: reduce the variance within classes and maximize the variance between classes**
  - unique classification, hard to compare between maps.
- **Python PySAL library implementation**
  - <https://pysal.org/mapclassify>

U.S. Census Bureau, 2000

## AVERAGE AGE, US COUNTIES

■ younger

■ older

classification method  
**QUANTILE**



classification method  
**STD. DEVIATION**



classification method  
**EQUAL INTERVAL**



20

average age of county residents

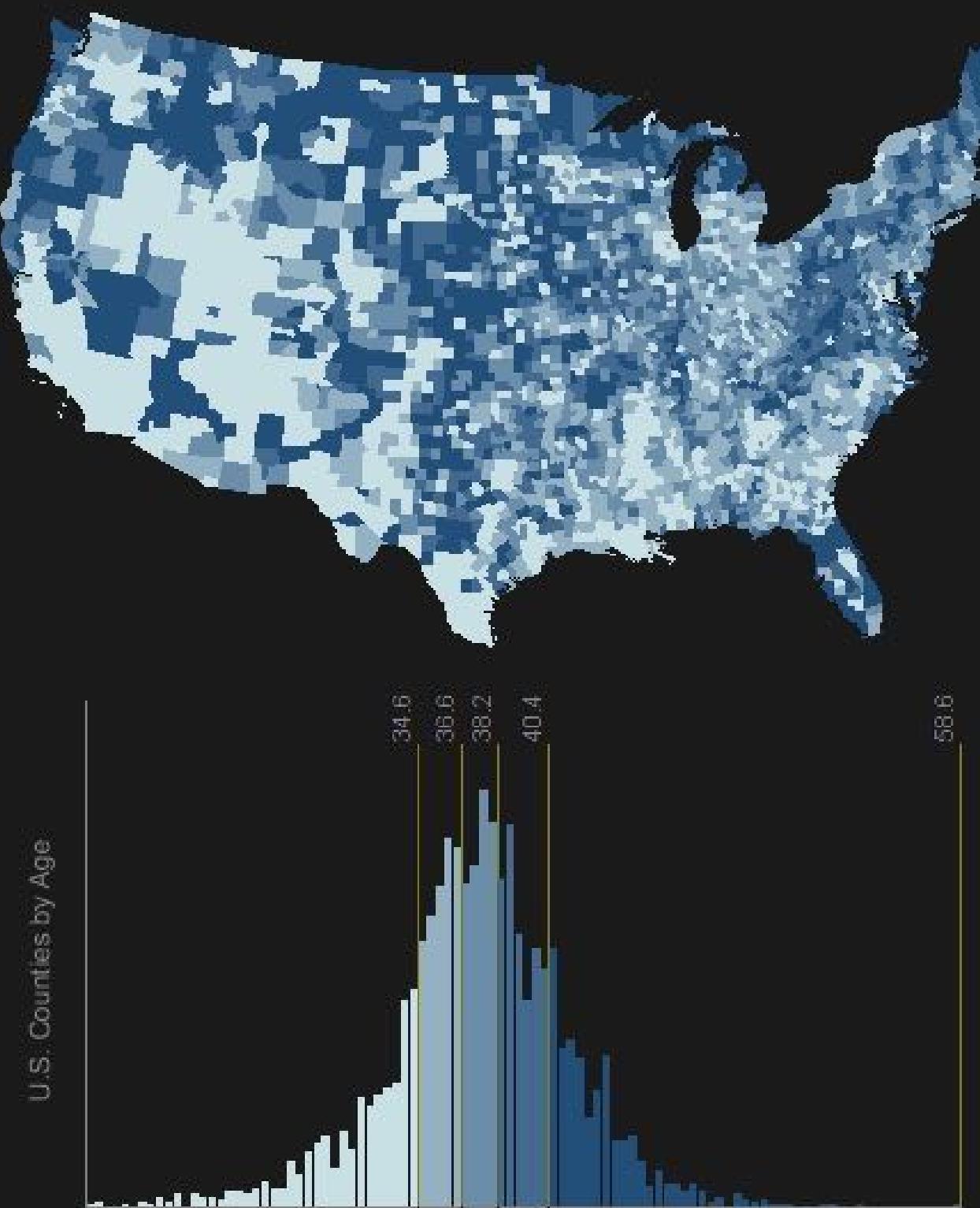
58

U.S. Census Bureau, 2000

## MEDIAN AGE

classification

## QUANTILE

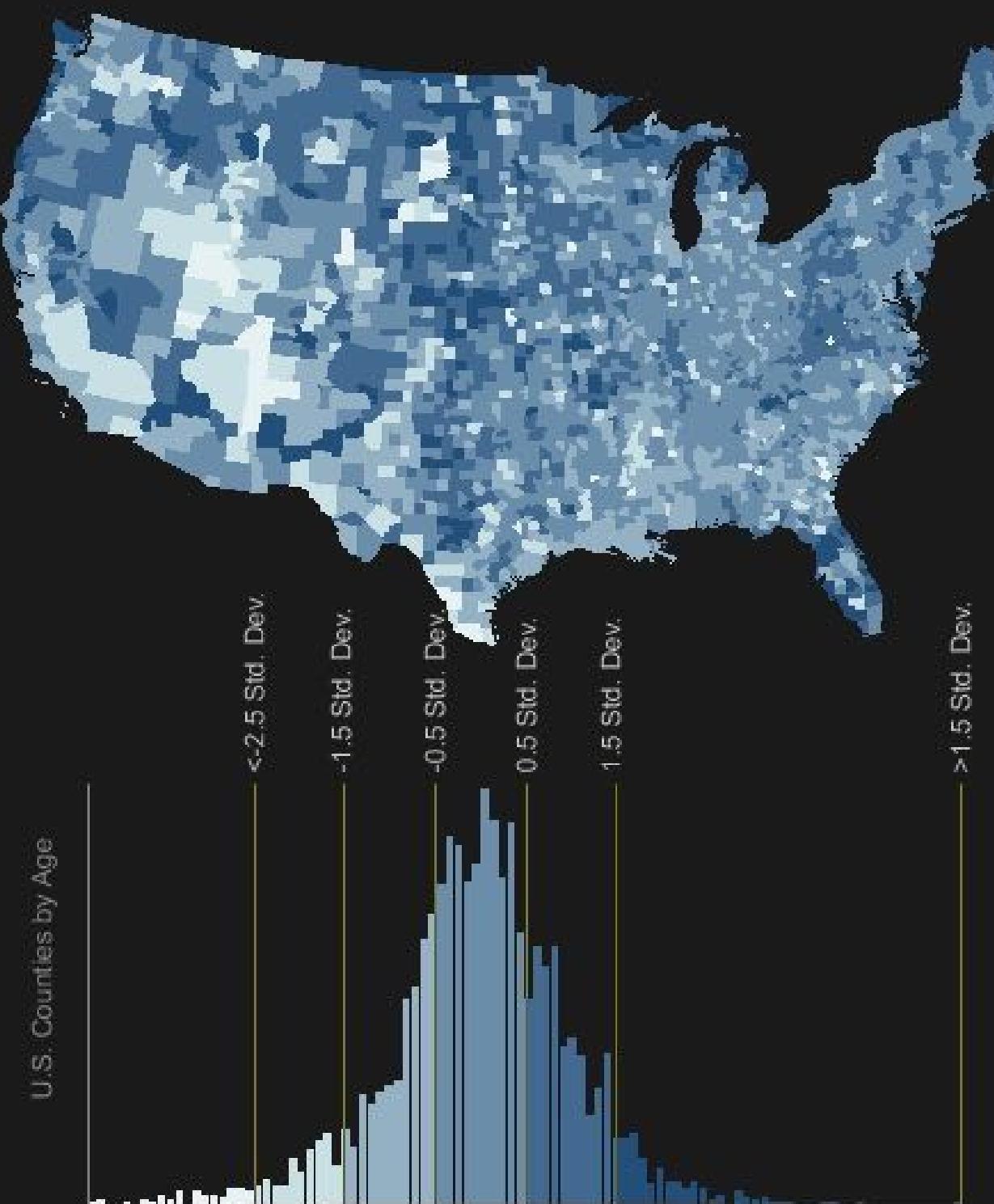


U.S. Census Bureau, 2000

## MEDIAN AGE

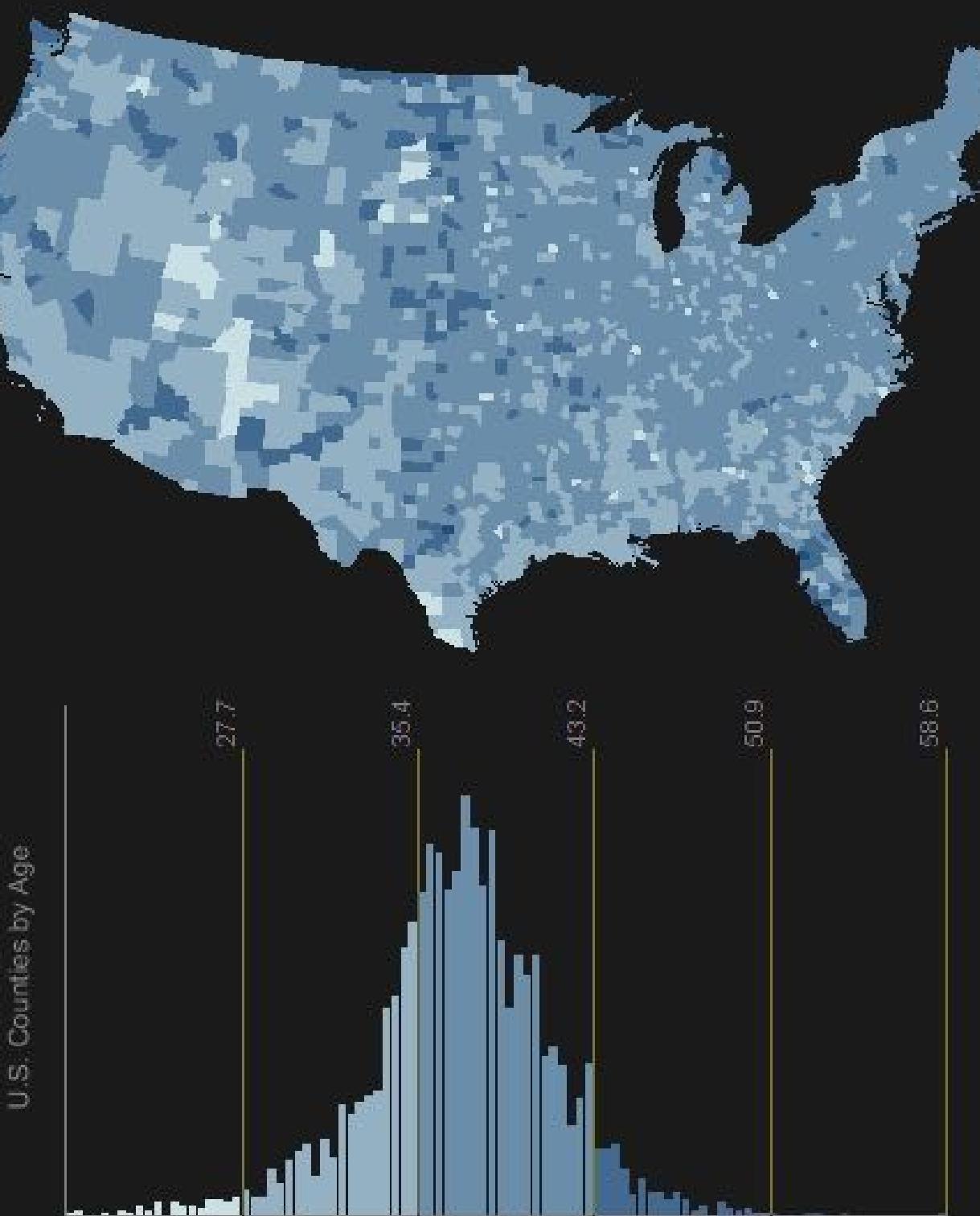
classification

## STD. DEVIATION

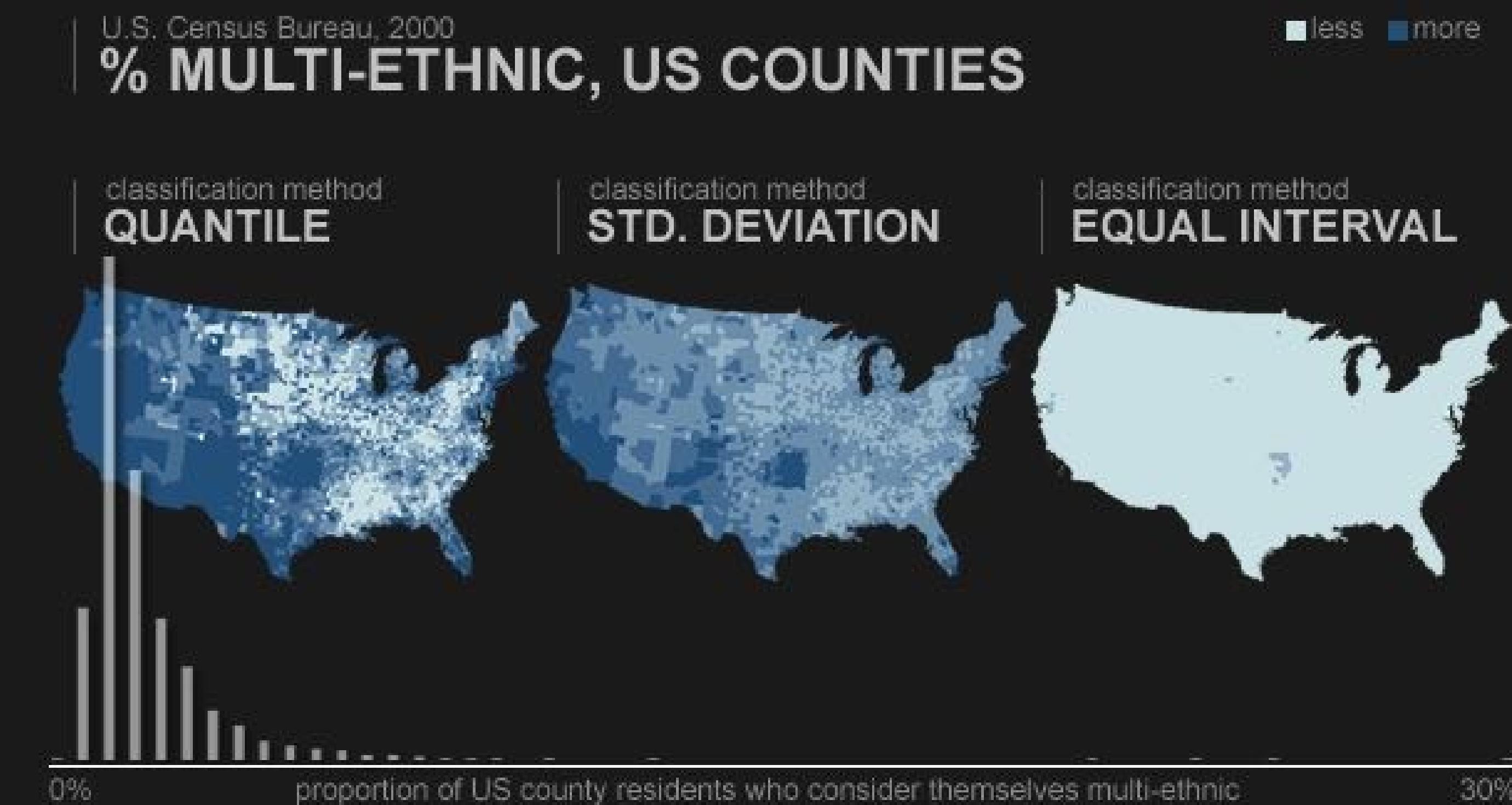


U.S. Census Bureau, 2000  
**MEDIAN AGE**

classification  
**EQUAL INTERVAL**



## Proportion of US county residents who consider themselves multiethnic

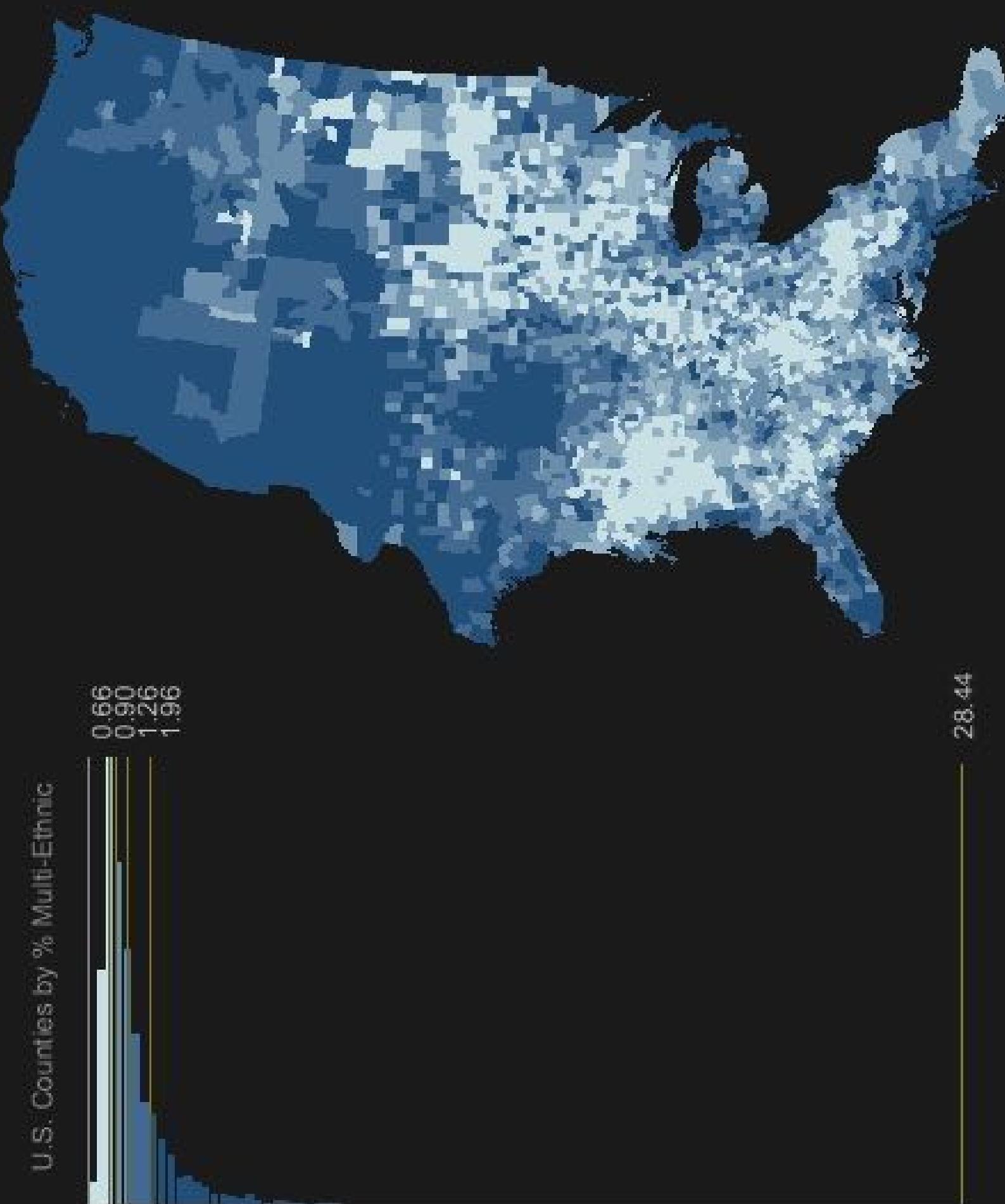


U.S. Census Bureau, 2000

## % MULTI-ETHNIC

classification

## QUANTILE

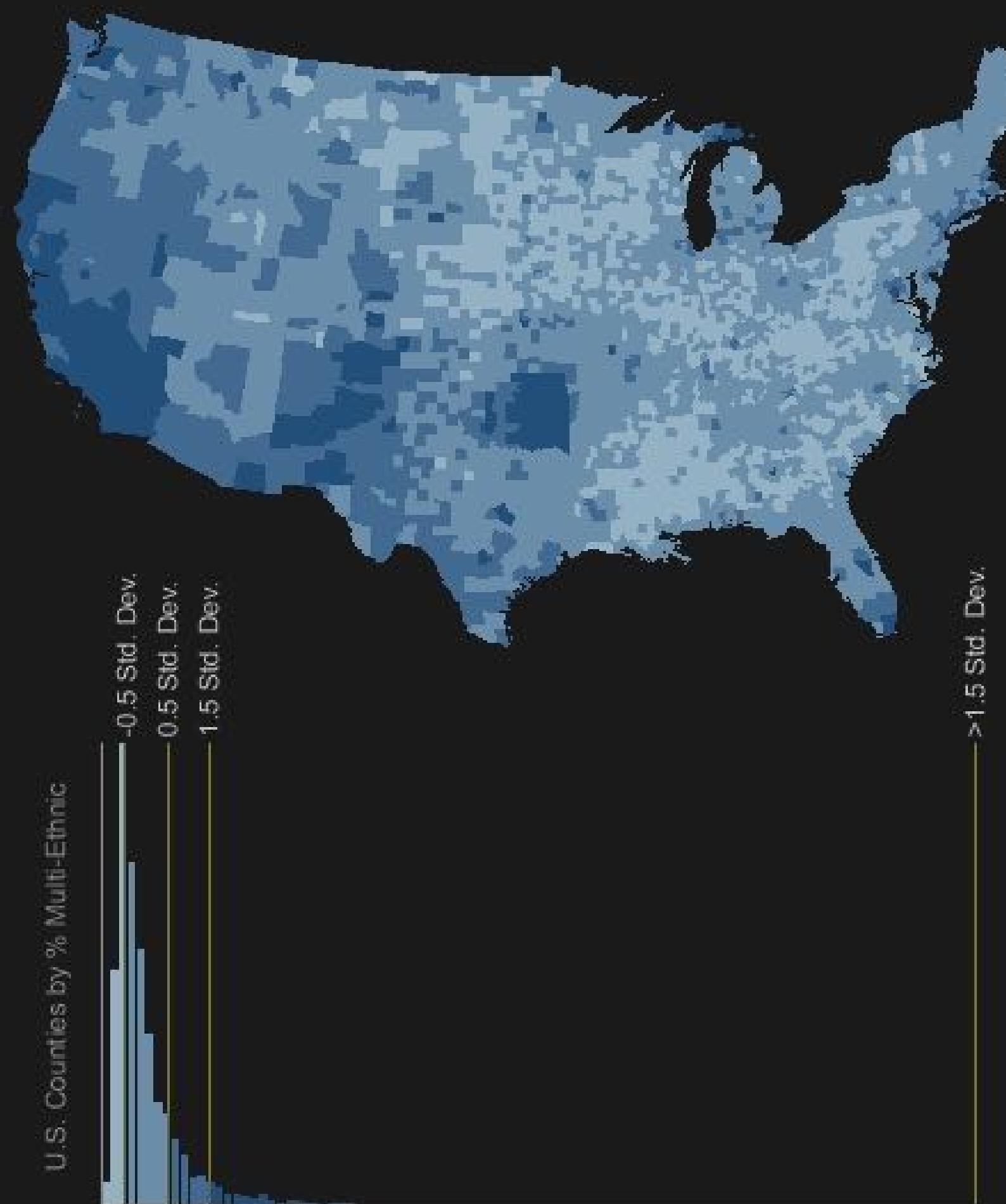


U.S. Census Bureau, 2000

## % MULTI-ETHNIC

classification

## STD. DEVIATION

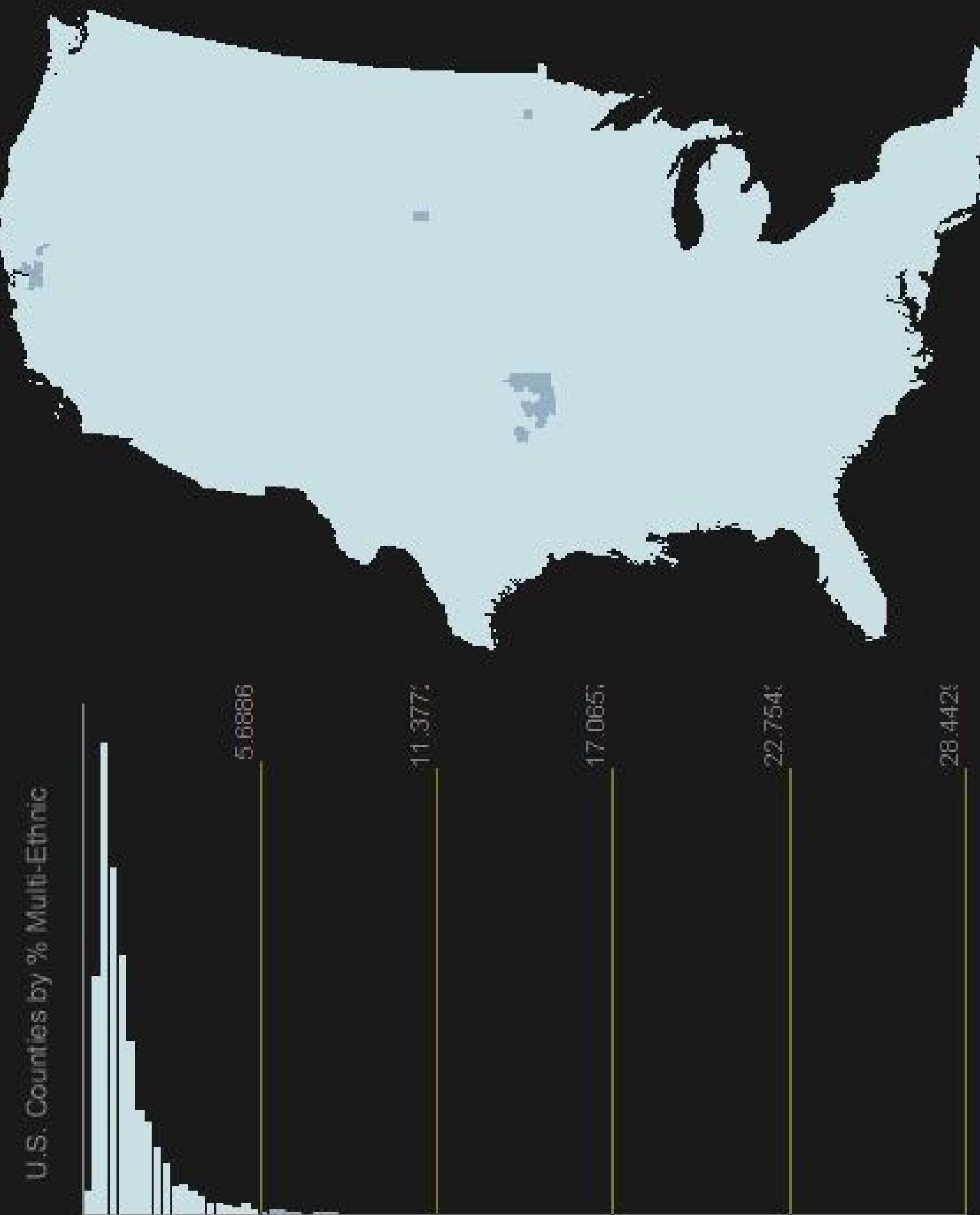


U.S. Census Bureau, 2000

## % MULTI-ETHNIC

classification

## EQUAL INTERVAL



# Symbolizing the choropleths

- **Select colors wisely!**
- **Monochrome shading**
  - darker is more!
  - vary density
  - different schemas: Munsell vs Stevens
- **Color shading**
  - hue is not always a good variable, unless bipolar distribution
  - use saturation or intensity

Number of data classes: 6

Nature of your data:  sequential  diverging  qualitative

Pick a color scheme:

Multi-hue:

Single hue:

Only show:

colorblind safe  print friendly  photocopy safe

Context:

roads  cities  borders

Background:

solid color  terrain

6-class BuGn

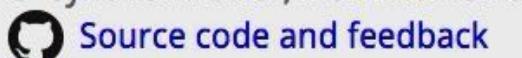
EXPORT

HEX

#edf8fb
#ccece6
#99d8c9
#66c2a4
#2ca25f
#006d2c

COLORBREWER 2.0  
color advice for cartography

© Cynthia Brewer, Mark Harrower and The Pennsylvania State University



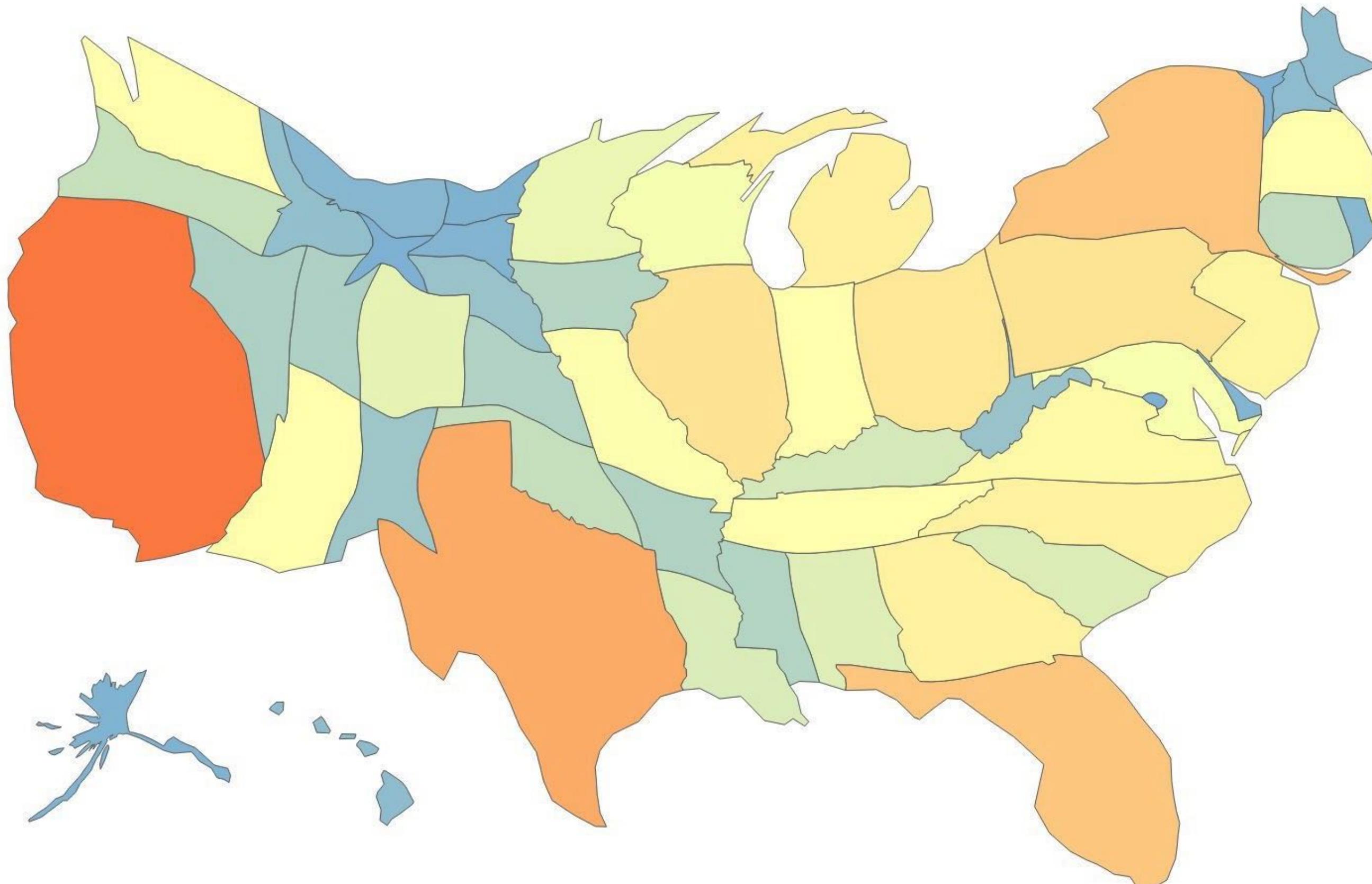
[Back to Flash version](#)

[Back to ColorBrewer 1.0](#)

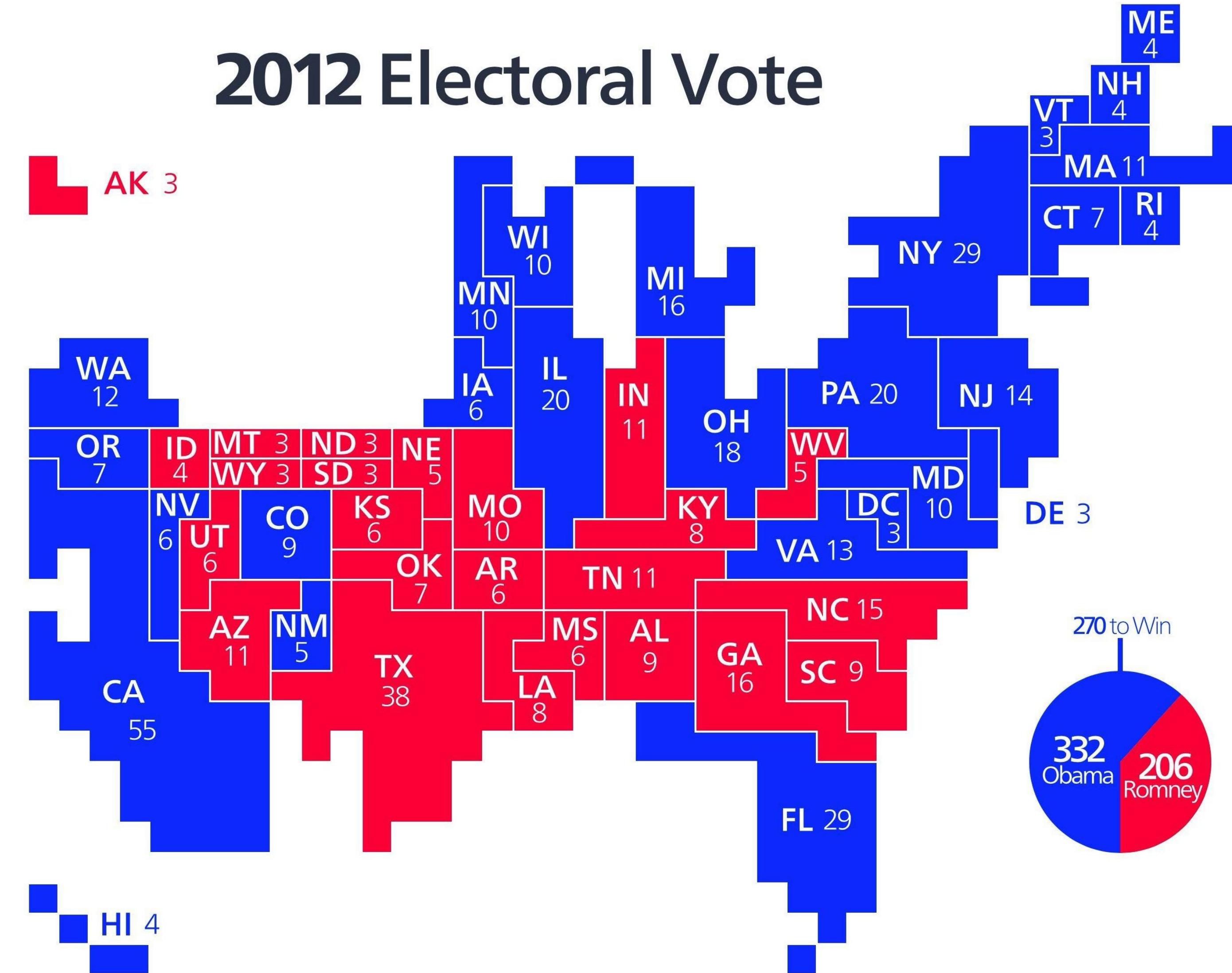


# cartograms

Scale by Population Estimate in 2010 calculated in 0.1 seconds



# 2012 Electoral Vote



# Flow Maps

- **Point pairs (one/two ways and symbol)    trajectories**
- **Encoding**
- **Edge between two locations indicates flow between those locations**
  - Width of edge proportional to flow
  - Usually wider end of edge is source of flow
- **Limitations**
  - Can get difficult to compare flows
  - Best flow maps are done by hand





Overview

## CORE LAYERS

[LineLayer](#)

HexagonLayer

IconLayer

GeoJsonLayer

ScreenGridLayer

ArcLayer

ScatterplotLayer

## CUSTOM LAYERS

Brushing Layer

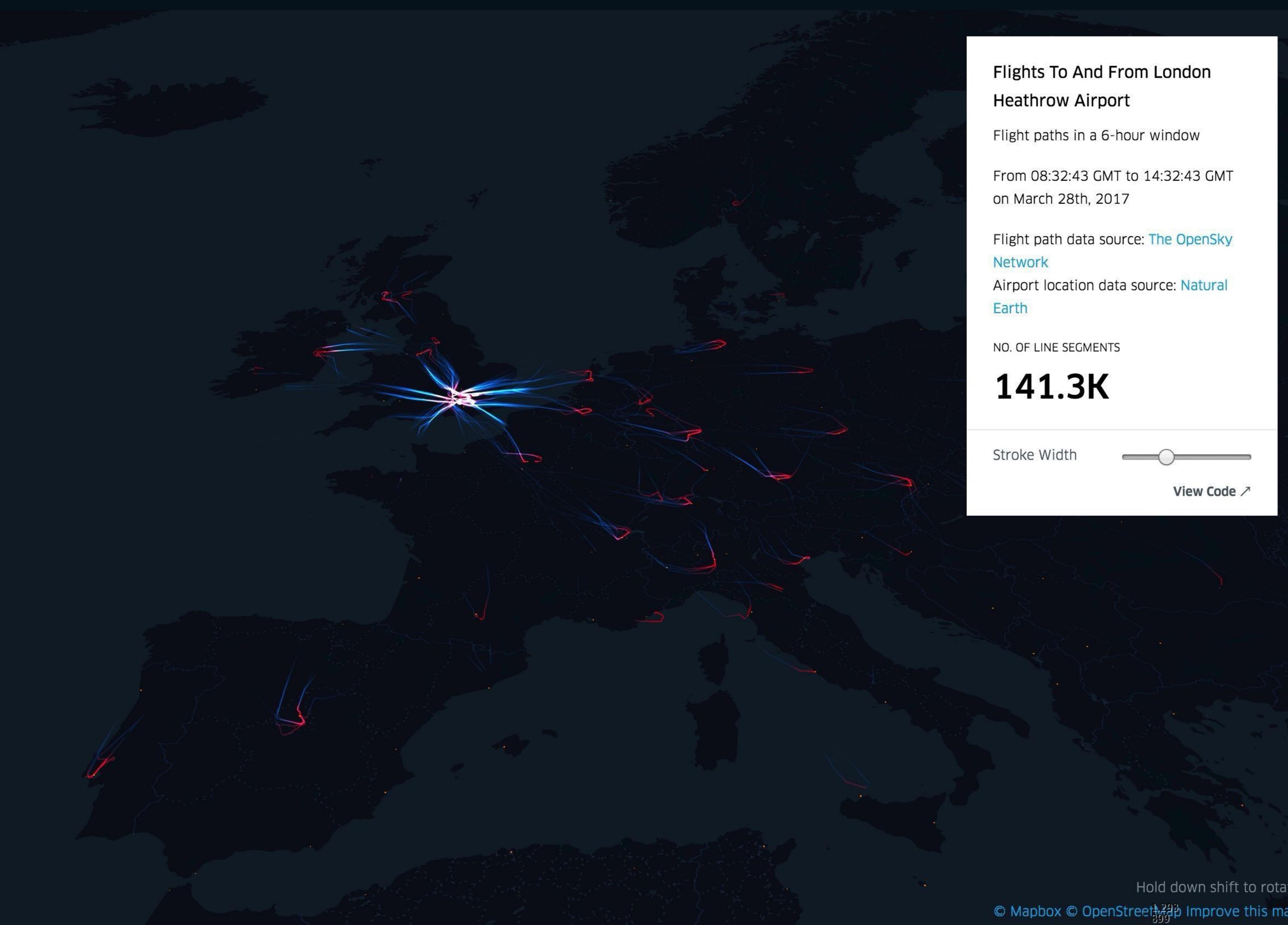
Trip Routes

Wind Map

## BEYOND MAPS

3D Surface Explorer

3D Indoor Scan



# Reading Material

[tutorials] [Geographical plotting in Geopandas](#) from the course repo

[user guide] [Official Geopandas User Guide](#)

# Questions?

:::::::



@rschifan



rossano.schifanella@unito.it



<http://www.di.unito.it/~schifane>