Introduction to Data Visualization

Visualizing Geospatial Data

https://worldle.teuteuf.fr/

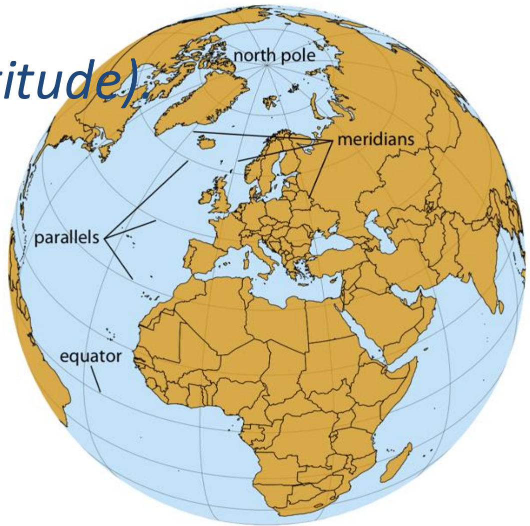
Halil Bisgin, Ph.D.

Geospatial Data-Maps

- Many datasets contain information linked to locations in the physical world.
  - *A dataset may contain information about where people with specific attributes (covid-19 cases, income, age, political affiliation, etc.)*
- It can be helpful to visualize the data in their proper geospatial context.
- Maps are intuitive, but they can be challenging to design.
  - *The choropleth map represents as differently colored spatial areas.*
  - *Cartograms, which may purposefully distort map areas or represent them in stylized form, for example as equal-sized squares.*

Projections

- The earth is an oblate spheroid.
- To uniquely specify a location on the earth, we need three pieces of information:
  - *where we are located along the direction of the equator (longitude),*
  - *how close we are to either pole when moving perpendicular to the equator (the latitude)*
  - *how far we are from the earth's center (the altitude)*
- Primarily: longitude and latitude



Longitude and latitude

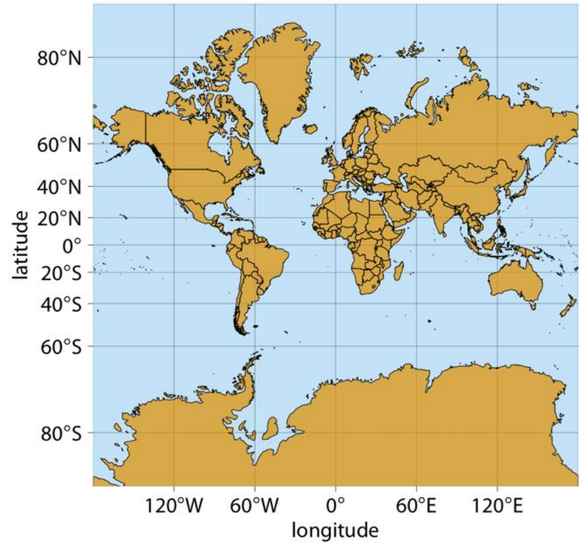
- They are both angles, expressed in degrees.
- Degrees longitude: how far east/west a location lies.
  - *Lines of equal longitude are referred to as meridians, which terminate at the two poles.*
- Degrees latitude: how far north or south a location lies.
  - *Lines of equal latitude are referred to as parallels, since they run parallel to the equator.*
- All meridians have the same length, corresponding to half of a great circle around the globe, whereas the length of parallels depends on their latitude

Projection

- To take the spherical surface of the earth and flatten it out so we can display it on a map.
- Can preserve either angles or areas but not both.
  - *A projection that does the former is called conformal and a projection that does the latter is called equal-area*
- Some projections preserve other quantities of interest, such as distances to some reference point.
- Others try to strike a compromise between preserving angles and areas
- Standards maintained by organizations:
  - *European Petroleum Survey Group (EPSG) and the Environmental Systems Research Institute (ESRI), maintain*

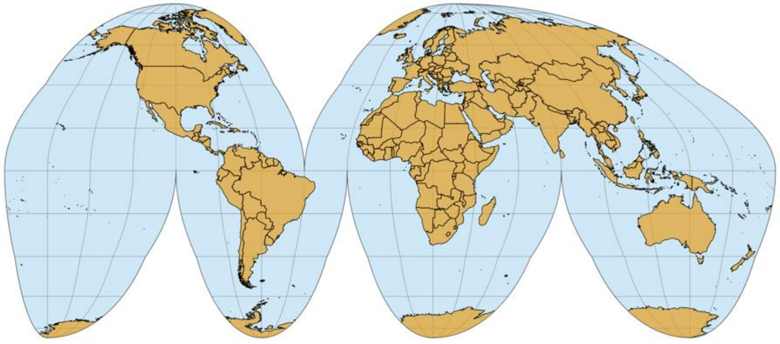
Mercator projection-Not favorite

- Conformal projection that accurately represents shapes but introduces severe area distortions near the poles.
  - *Meridians are evenly spaced vertical lines*
  - *Parallels are horizontal lines whose spacing increases the further we move away from the equator to make meridians evenly spaced*



Goode homolosine

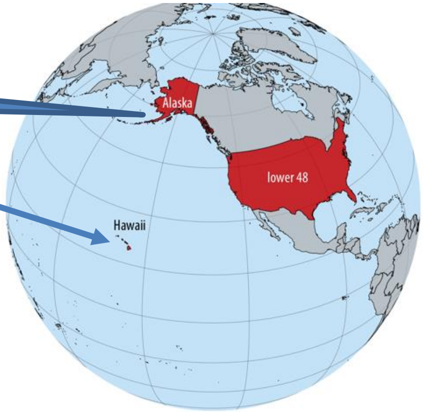
- Whole-world projection that is perfectly area-preserving
- Interrupted form:
  - *one cut in the northern hemisphere and three cuts in the southern hemisphere which don't interrupt major land masses.*
- The cuts allow the projection to both preserve areas and approximately preserve angles,
  - *at the cost of noncontiguous oceans, a cut through the middle of Greenland*
  - *several cuts through Antarctica.*



Distortions are at all levels

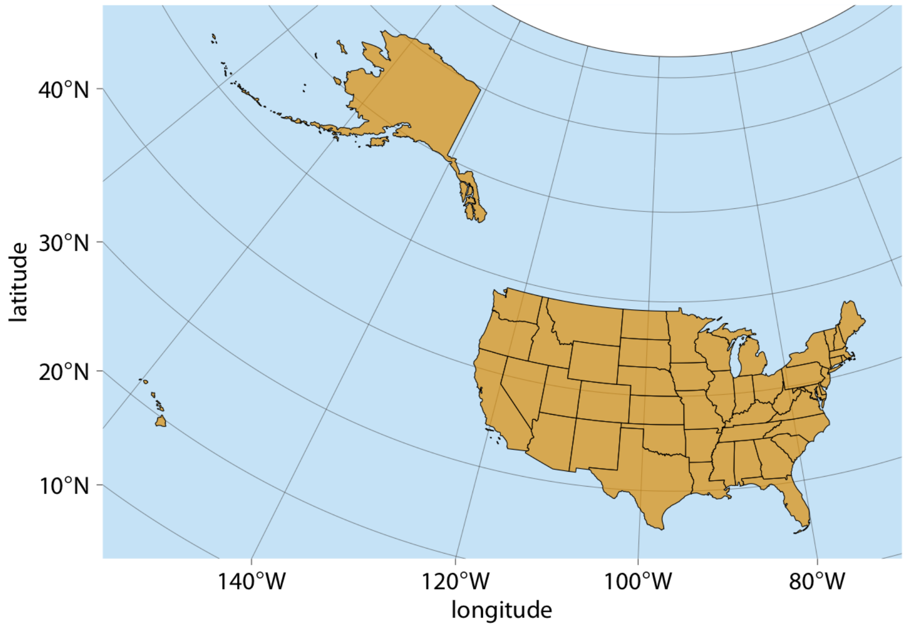
- Shape or area distortions due to map projections are particularly prominent when we're attempting to make a map of the whole world
- They can cause trouble even at the scale of individual continents or countries.
  - *United States, which consists of the lower 48 (which are 48 contiguous states), Alaska, and Hawaii*

- 48 alone are reasonably easy to project
- Alaska and Hawaii are so distant



Problems

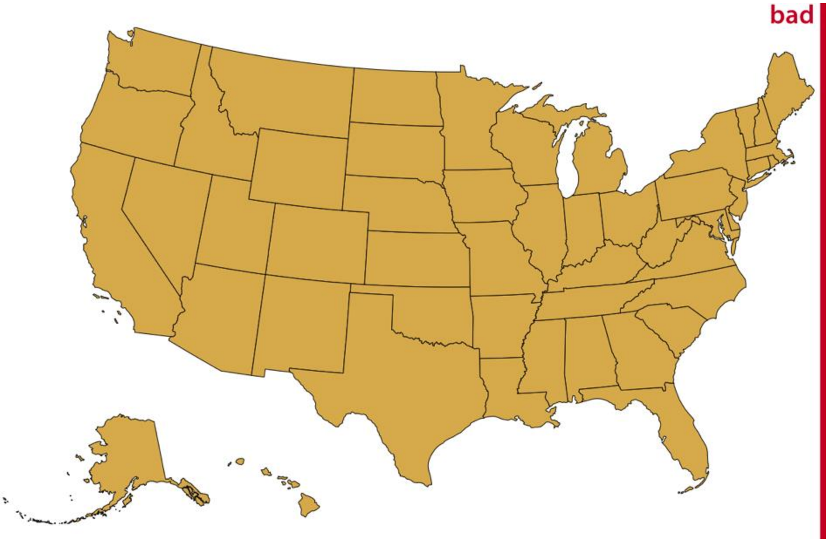
- (Equal-area) Albers projection provides a reasonable representation of the relative shapes, areas, and locations of the 50 states, but still issues.
  - *Alaska seems weirdly stretched out.*
  - *Dominated by ocean/empty space*



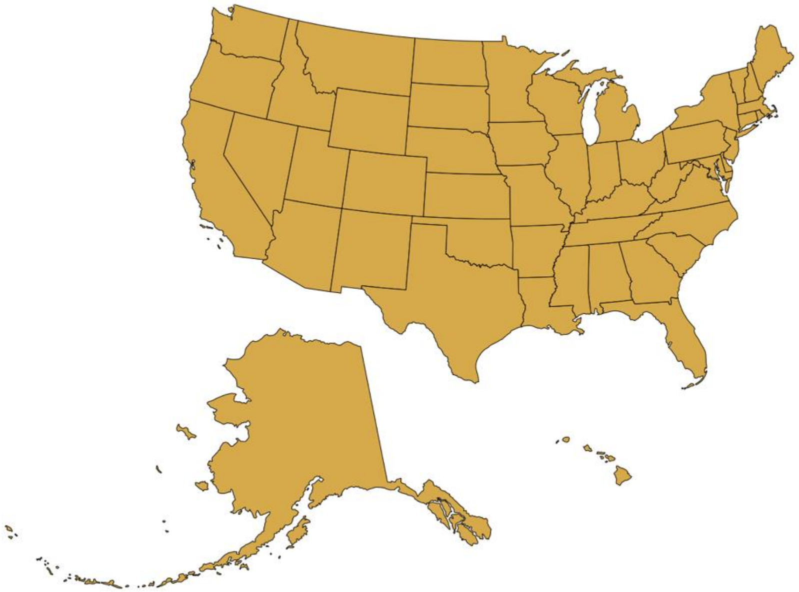
Some improvements

- We can project Alaska and Hawaii separately.

Alaska looks much smaller (misleading)

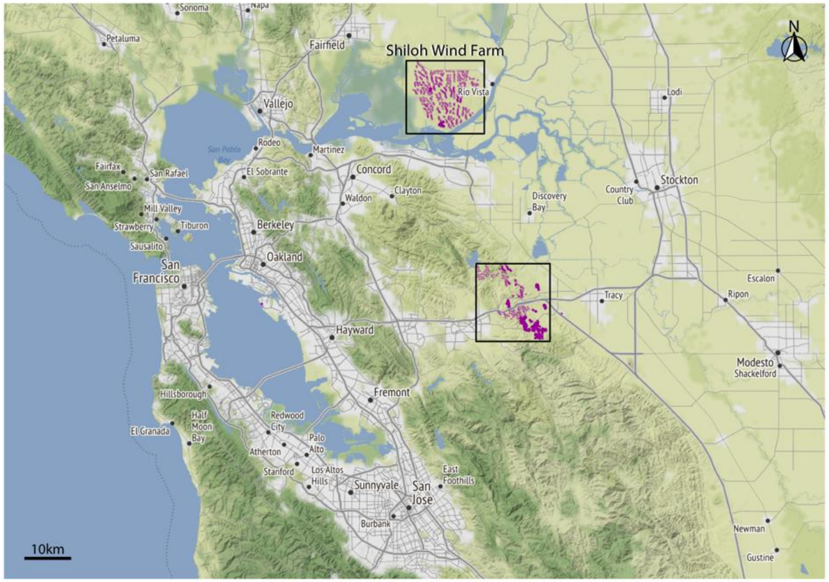


Reveals that Alaska is the largest state



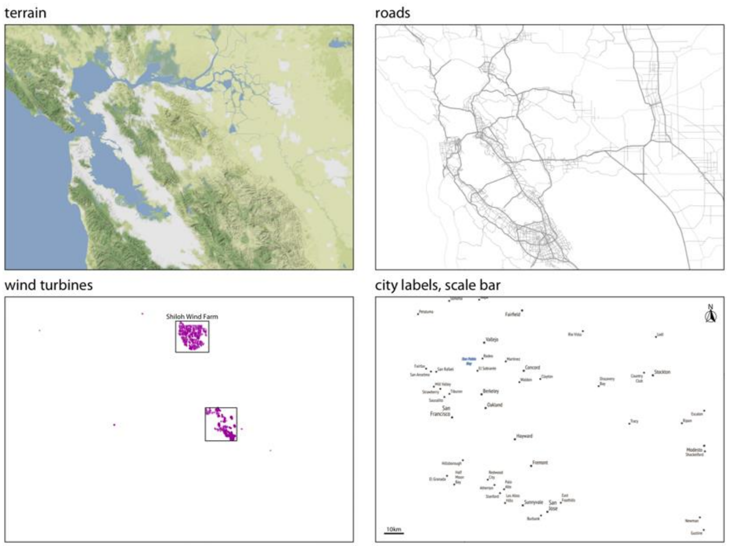
Layers

- To visualize geospatial data in the proper context, we usually create maps consisting of multiple layers showing different types of information.
  - *Wind turbines in the San Francisco Bay area*



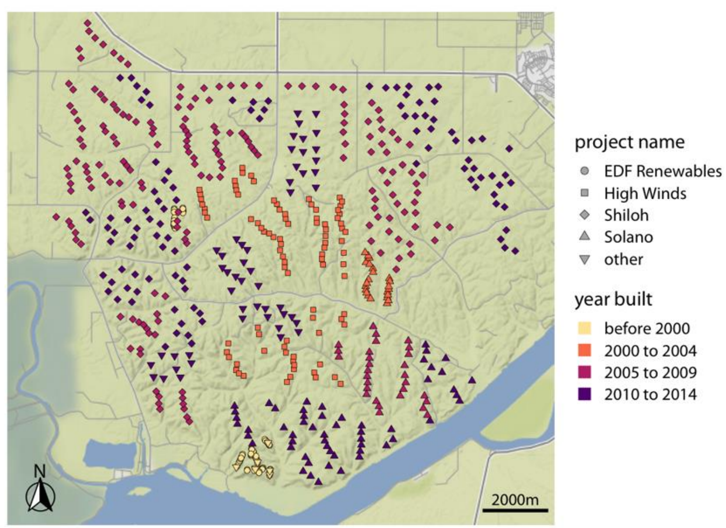
Layers

1. The terrain layer for hills, valleys, and water
2. The road network
3. The locations of individual wind turbines.
4. The locations and names of cities.



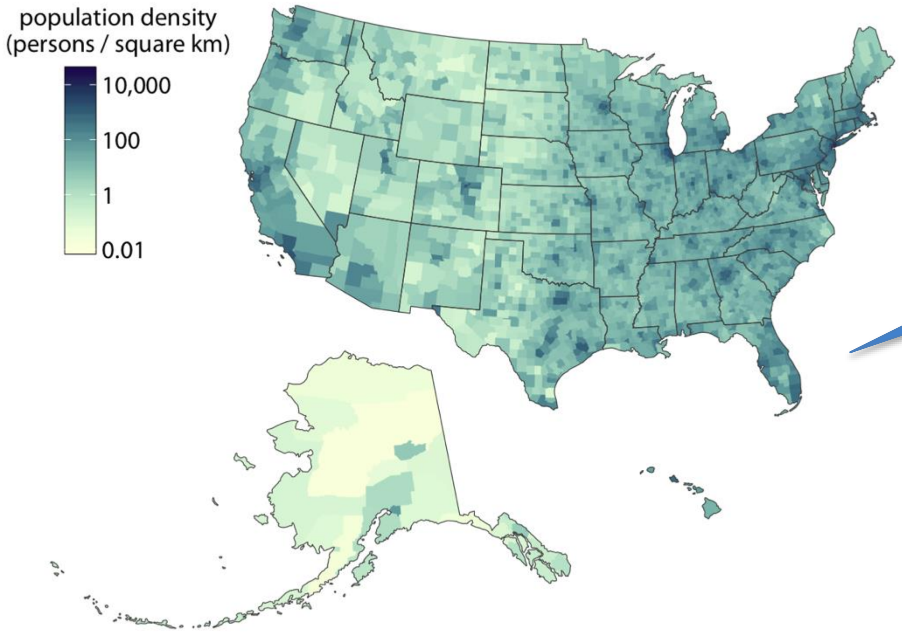
Mapping data onto aesthetics in maps

- We can place data points into their geographic context and show other data dimensions via aesthetics such as color or shape
  - *Individual wind turbines are shown as dots,*
  - *Color: when a specific turbine was built*
  - *Shape: the project to which the wind turbine belongs*



Choropleth Mapping

- Choropleth maps: Coloring individual regions in a map according to the data dimension we want to display.



Population density for US counties.

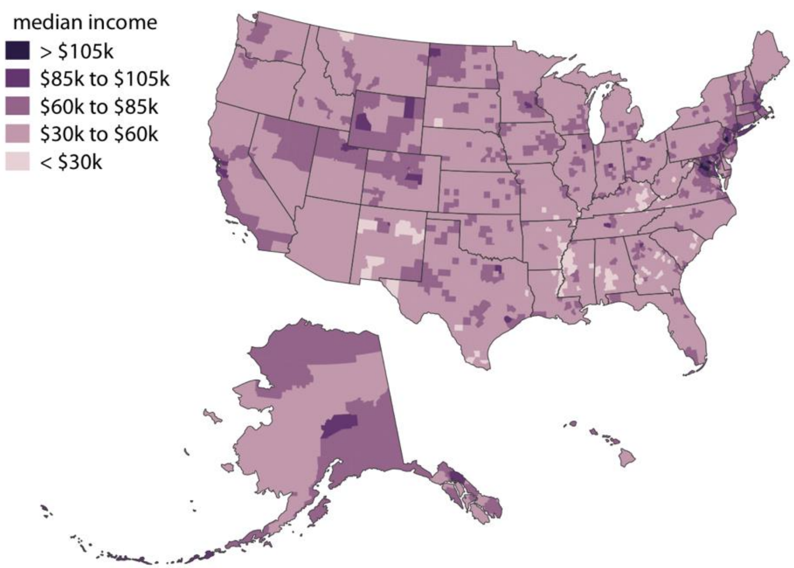
Notes on Choropleth

- Works best when the coloring represents a density (i.e., some quantity divided by surface area)
- We often see choropleths colored according to some quantity that is not a density.
  - *Needs caution as the size may convey unintended message*

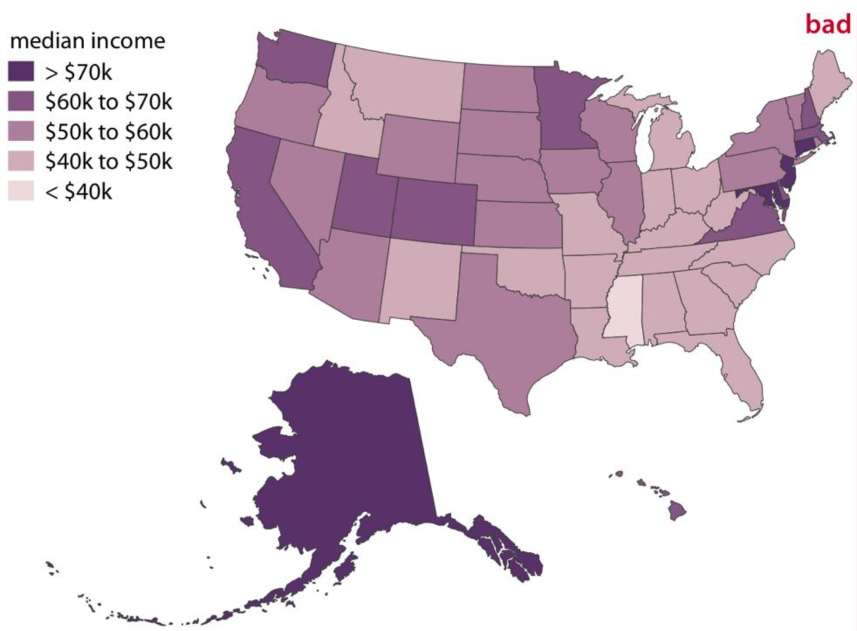
Continuous vs. discrete color scales

- While continuous color scales tend to look visually appealing, they can be difficult to read.

County level

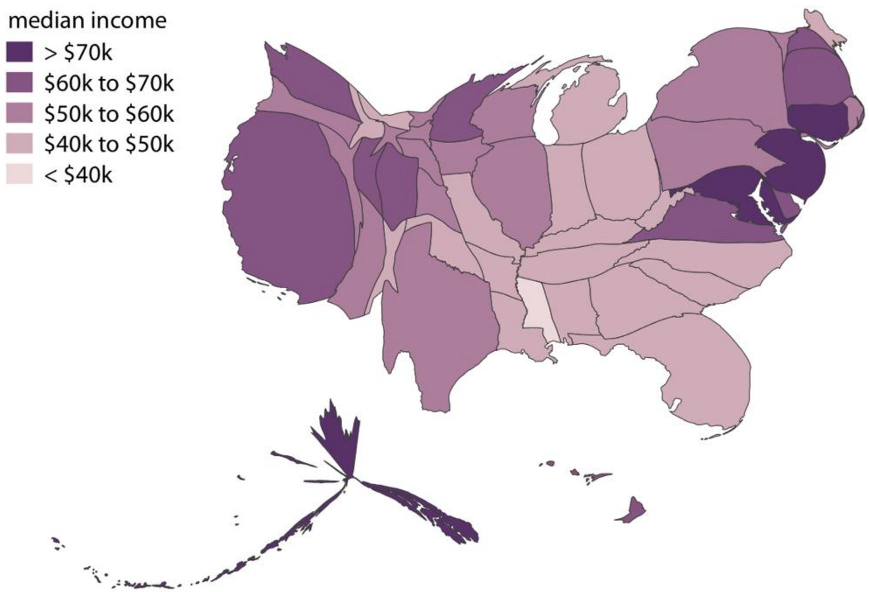


State level



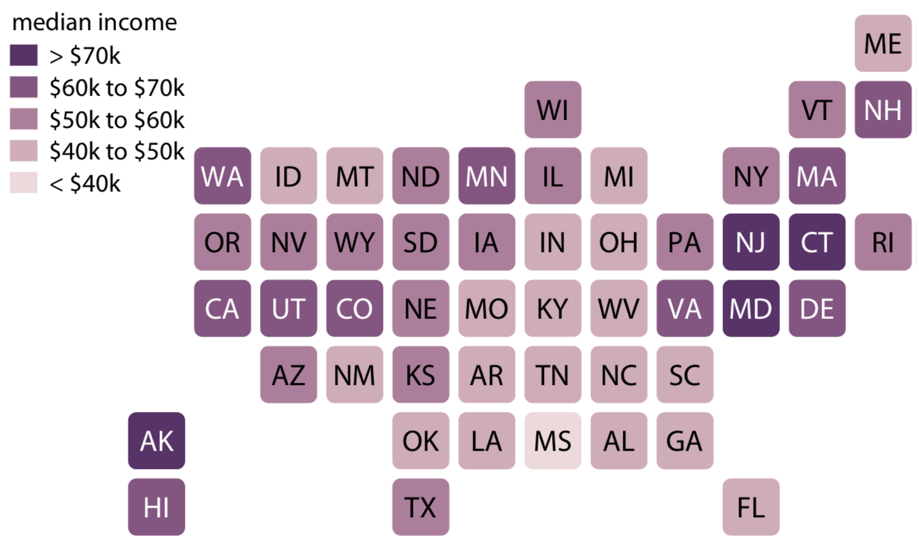
Cartograms

- Not every map-like visualization has to be geographically accurate to be useful.
- What if we deformed the states so their size was proportional to their number of inhabitants?



Cartogram heatmap

- E.g., each state is represented by a colored square.
  - *Does not correct for the population number in each state, and thus underrepresents (overrepresents) more (less) populous states*
  - *Treats all equally and doesn't weight them arbitrarily by their shape or size.*



Ex. Unemployment rate over time

- Draw an individual graph for each state
- Arrange based on the approximate relative positions
  - *If you're familiar with US map, better than alphabetical order*