Chapter 15

Fundamentals of Data Visualization

May 4, 2023

Visualizing geospatial data

- Many datasets contain information linked to locations:
  - where specific plants or animals have been found
  - where people with specific attributes (such as income, age, or educational attainment) live
  - where man-made objects (e.g., bridges, roads, buildings) have been constructed
- Maps tend to be intuitive to readers but they can be challenging to design.
- The choropleth map represents data values as differently colored spatial areas.
- Cartograms purposefully distort map areas or represent them in stylized form, for example as equal-sized squares.

World coordinates

- To specify a place:
  latitude, longitude, altitude
- A reference system for these is called a **datum**
- One widely used datum is the World Geodetic System (WGS) 84, which is used by the Global Positioning System (GPS).

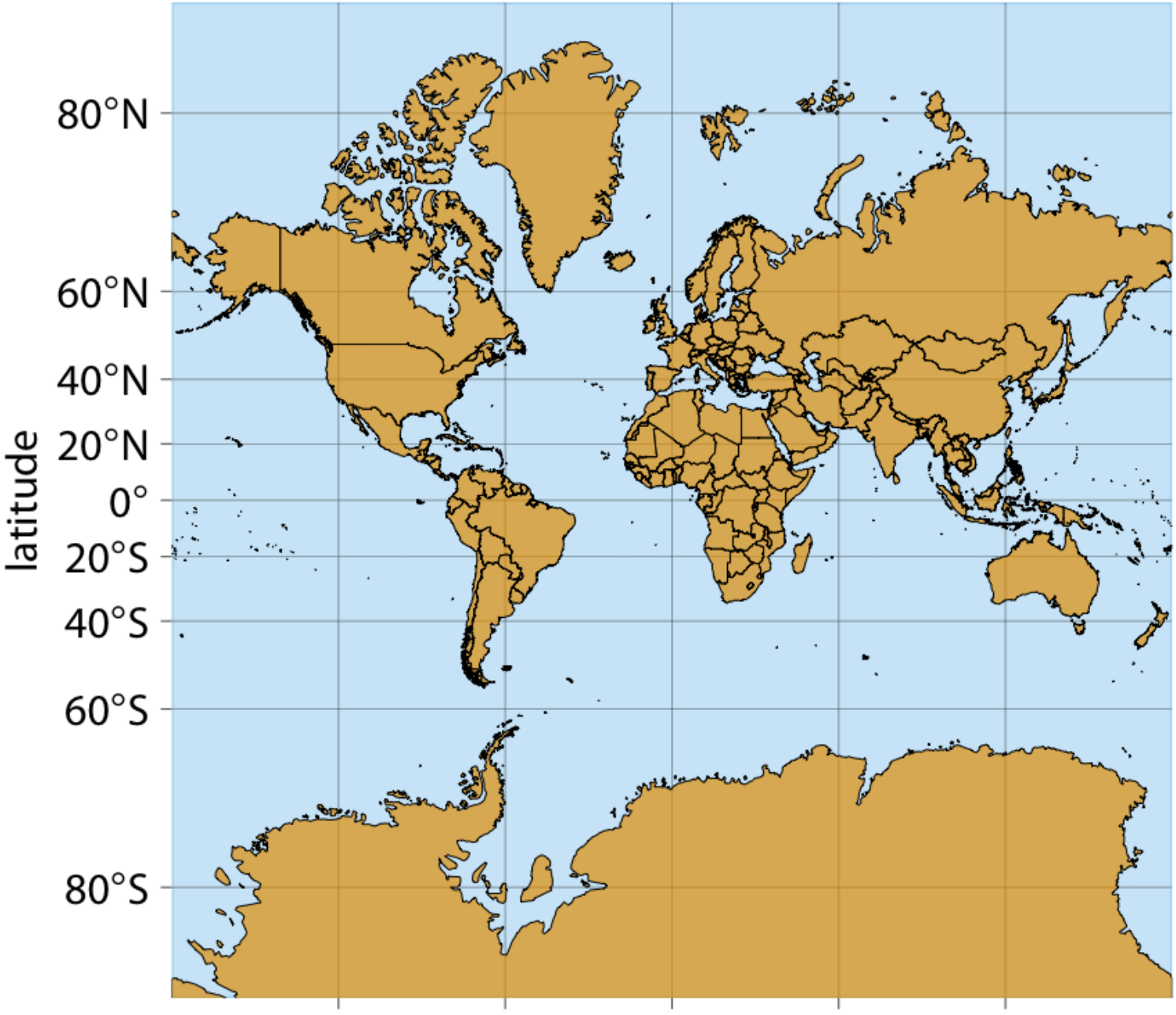
World coordinates

- Frequently altitude is not recorded.
- Lines of equal longitude are **meridians**
- Lines of equal latitude are called **parallels**
- The **prime meridian** is at 0° longitude
- The **equator** is at 0° latitude

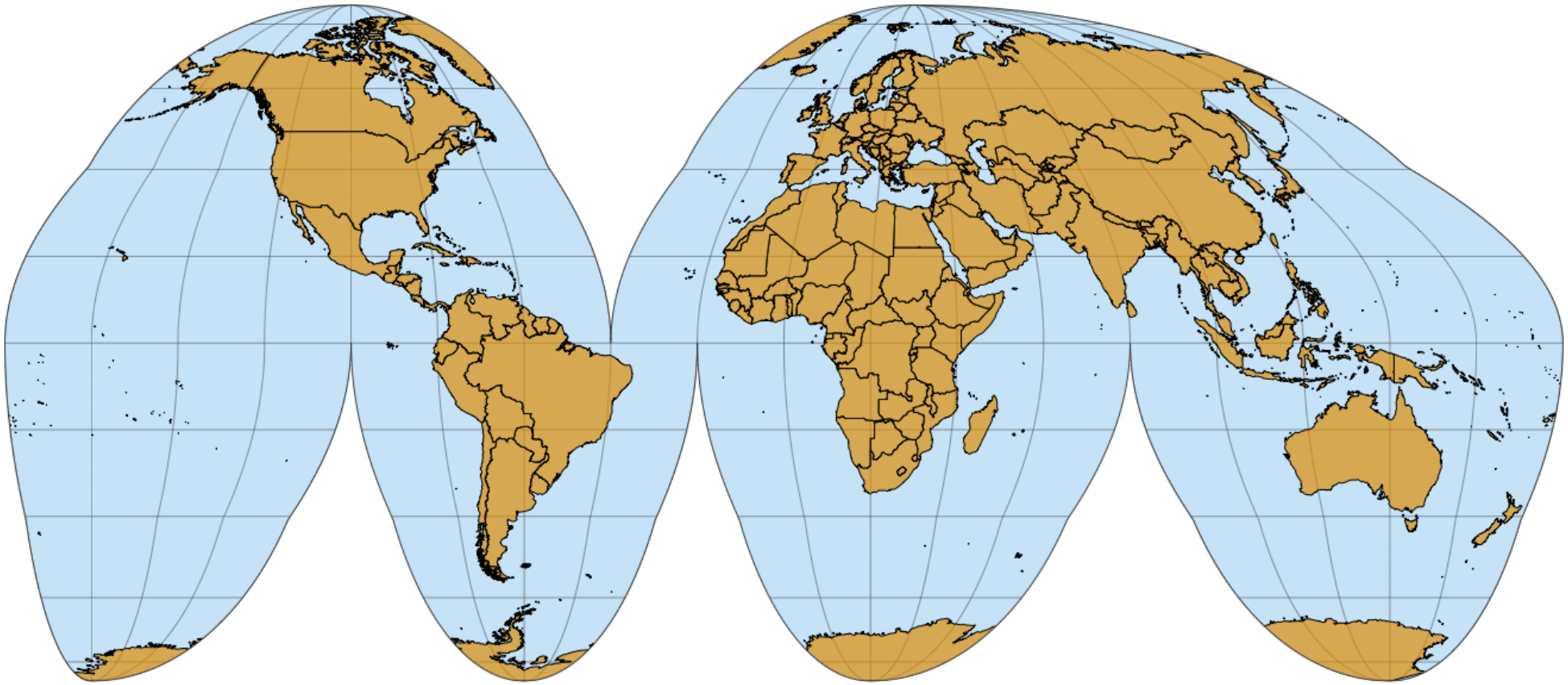
Projections

- In map-making we need to take the spherical surface of the earth and flatten it out.
- This projection introduces distortions.
- The projection can preserve either angles or areas.
- A projection that preserves angles is called **conformal**.
- A projection that preserves areas is called **equal-area**.
- Other projections may instead preserve other quantities of interest, such as distances to some reference point or line.
- Some projections attempt to strike a compromise between preserving angles and areas.
- These compromise projections are frequently used to display the entire world in an aesthetically pleasing manner.

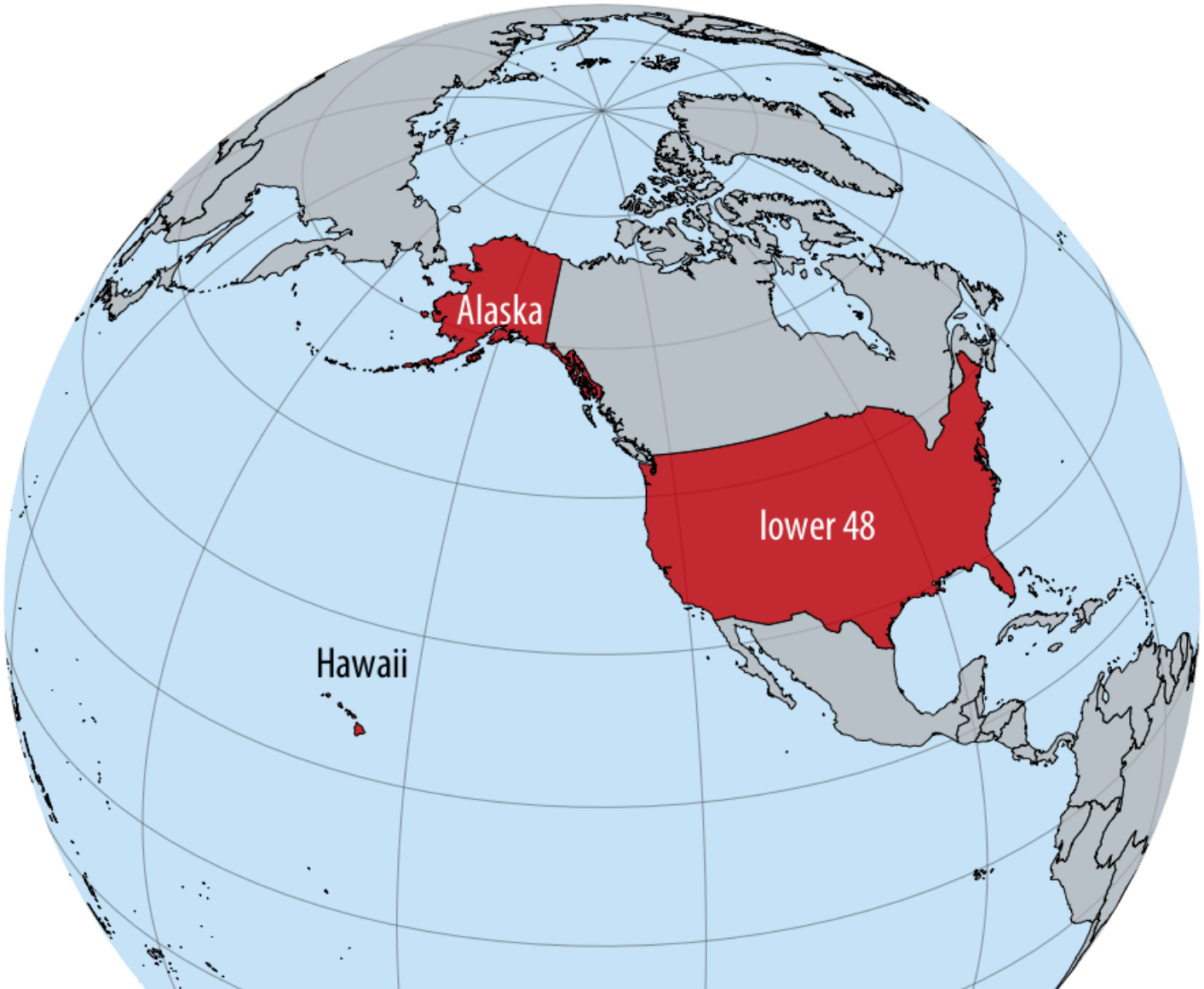
Mercator: a conformal projection



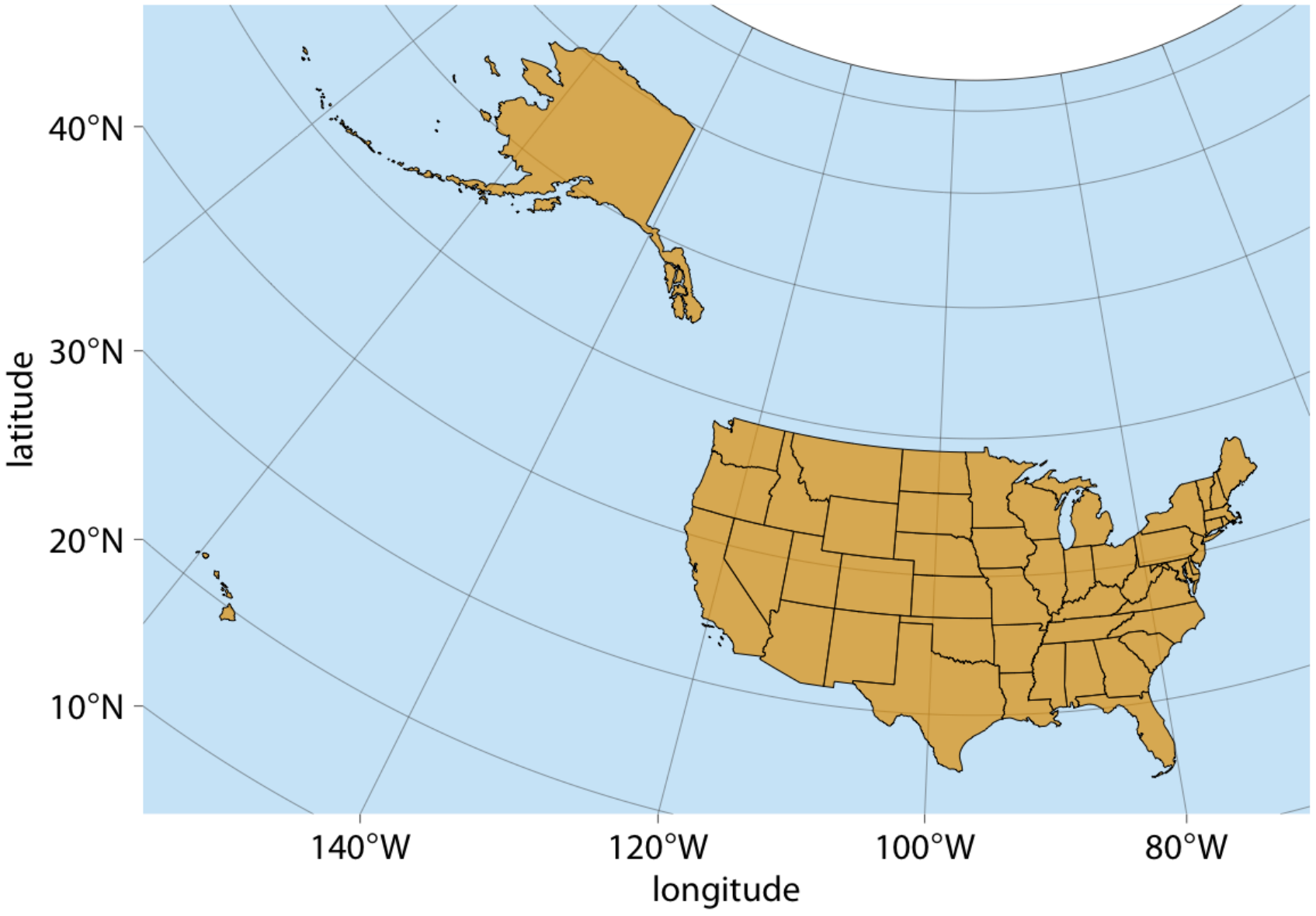
Goode homolosine: an equal-area projection



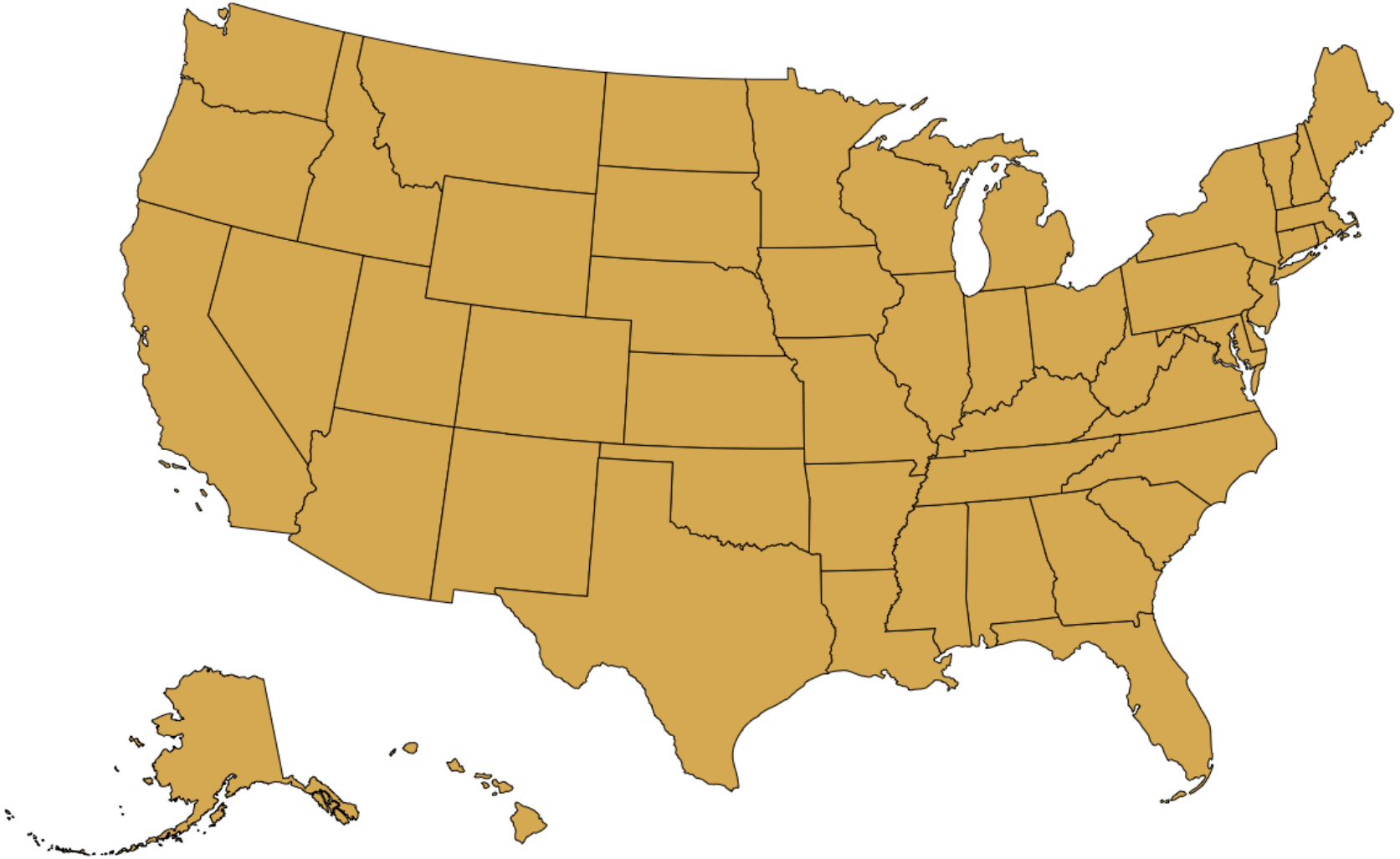
A challenge: the USA



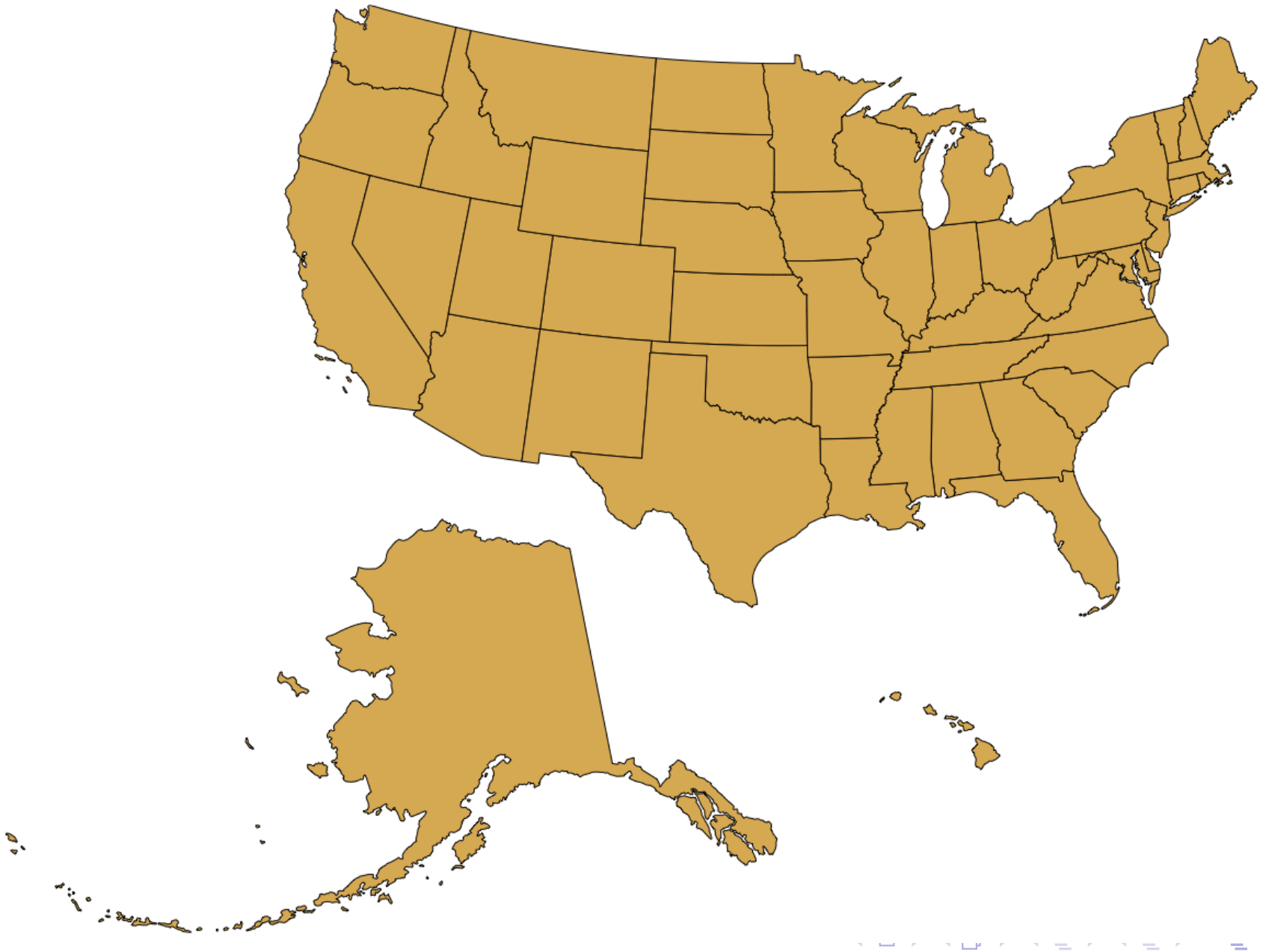
Equal-area Albers projection



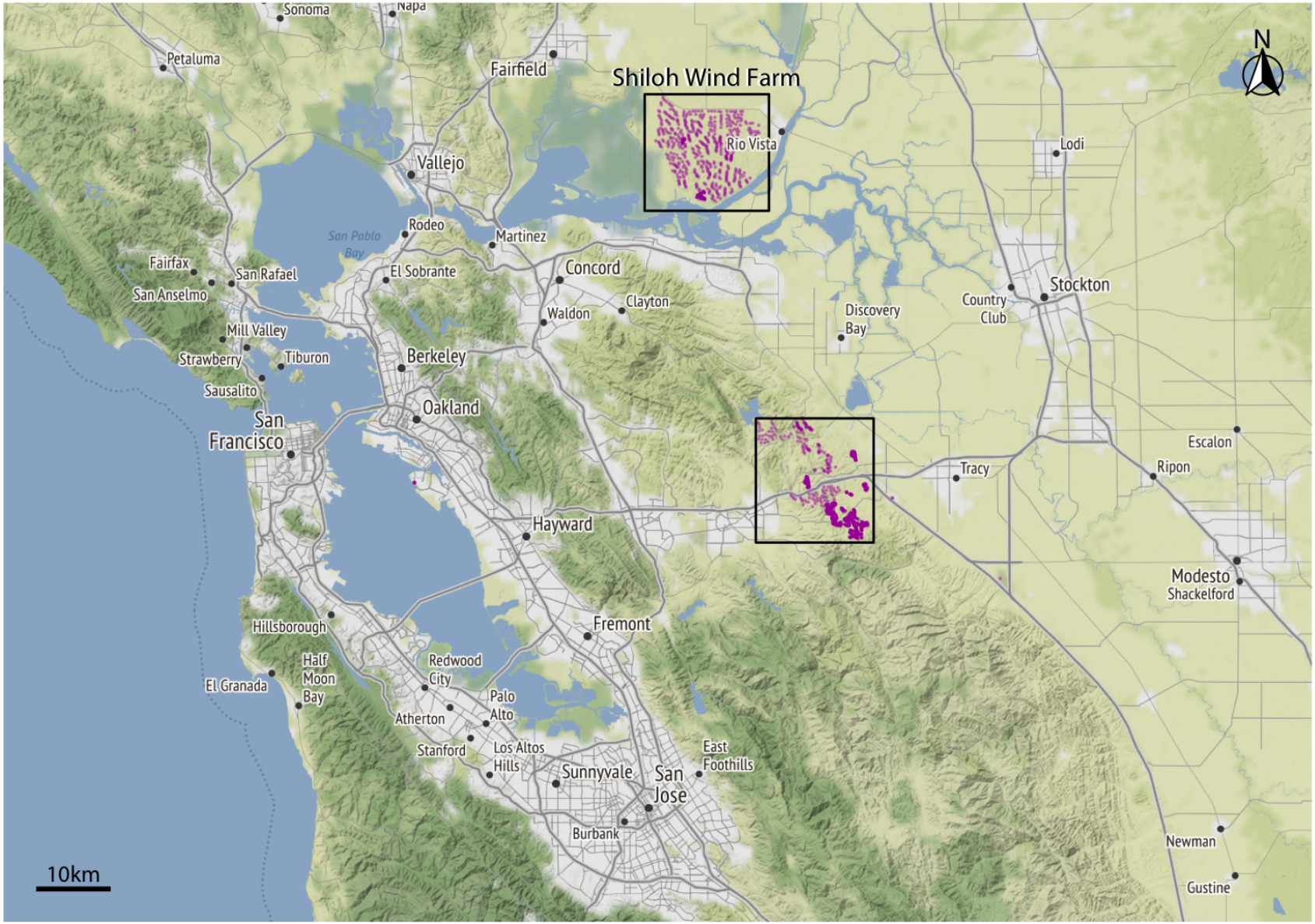
Common practice



Preserving areas



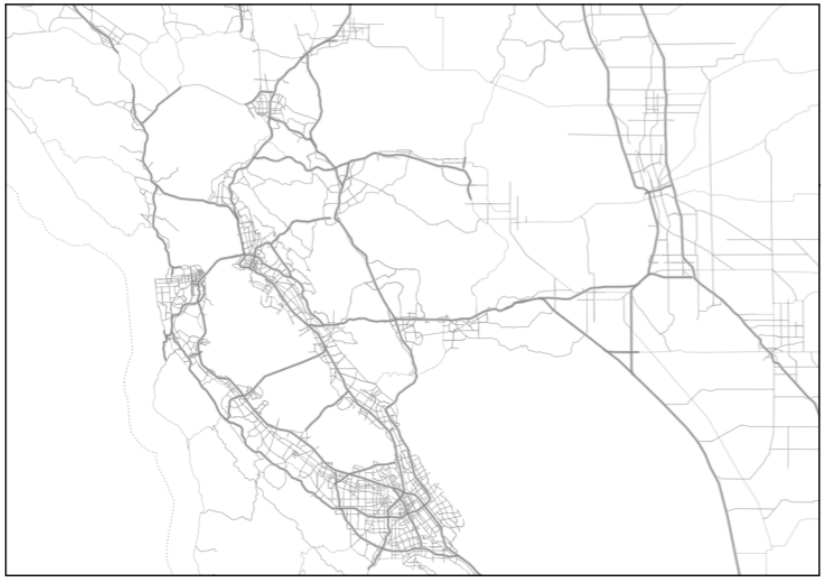
Layers


Shiloh Wind Farm
N
Sonoma
Napa
Petaluma
Fairfield
Rio Vista
Lodi
Vallejo
Rodeo
Martinez
San Pablo Bay
Fairfax
San Rafael
El Sobrante
Concord
San Anselmo
Clayton
Waldon
Stockton
Country Club
Discovery Bay
Mill Valley
Strawberry
Tiburon
Berkeley
Sausalito
Oakland
San Francisco
Escalon
Tracy
Ripon
Hayward
Modesto
Shackelford
Hillsborough
Fremont
Half Moon Bay
El Granada
Redwood City
Palo Alto
Atherton
Stanford
Los Altos Hills
East Foothills
Sunnyvale
San Jose
Burbank
Newman
Gustine
10km

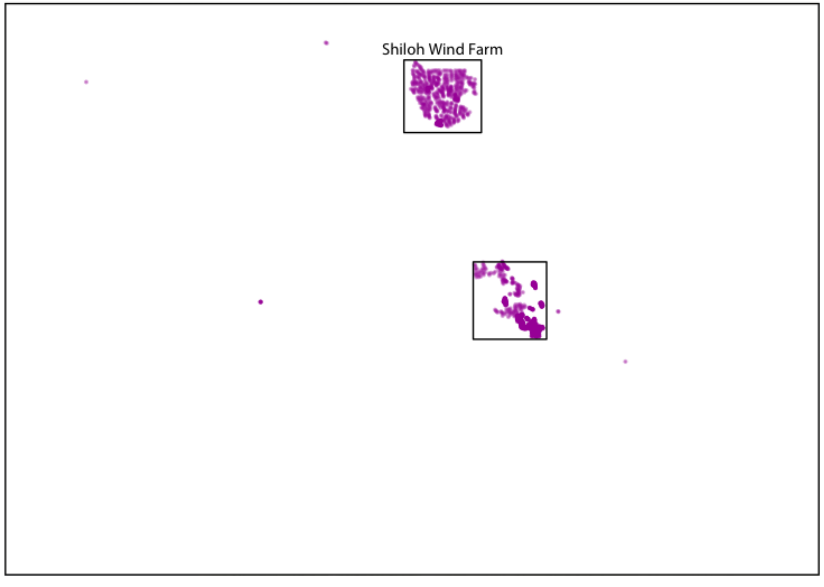
Layers

terrain

roads

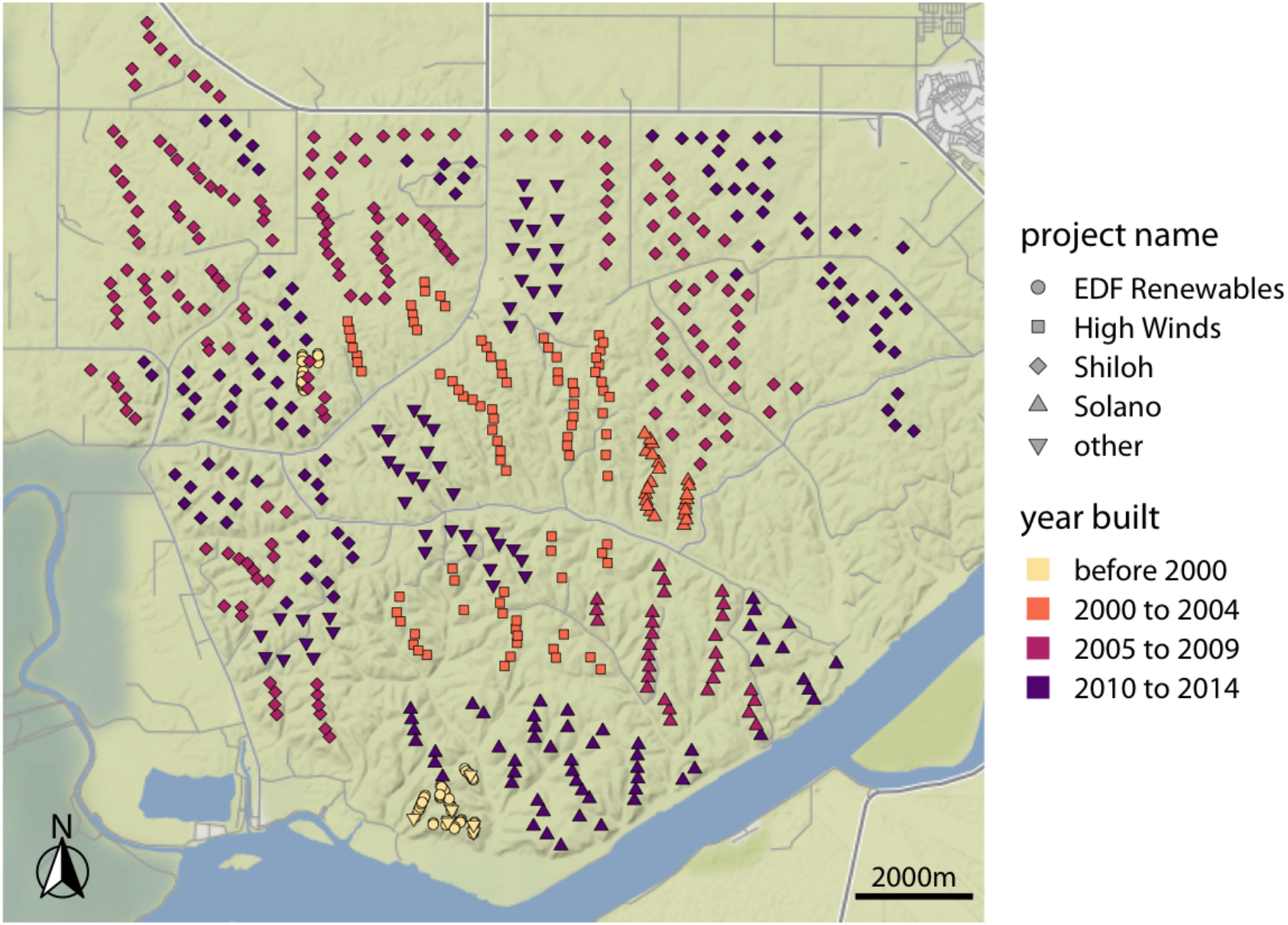


wind turbines

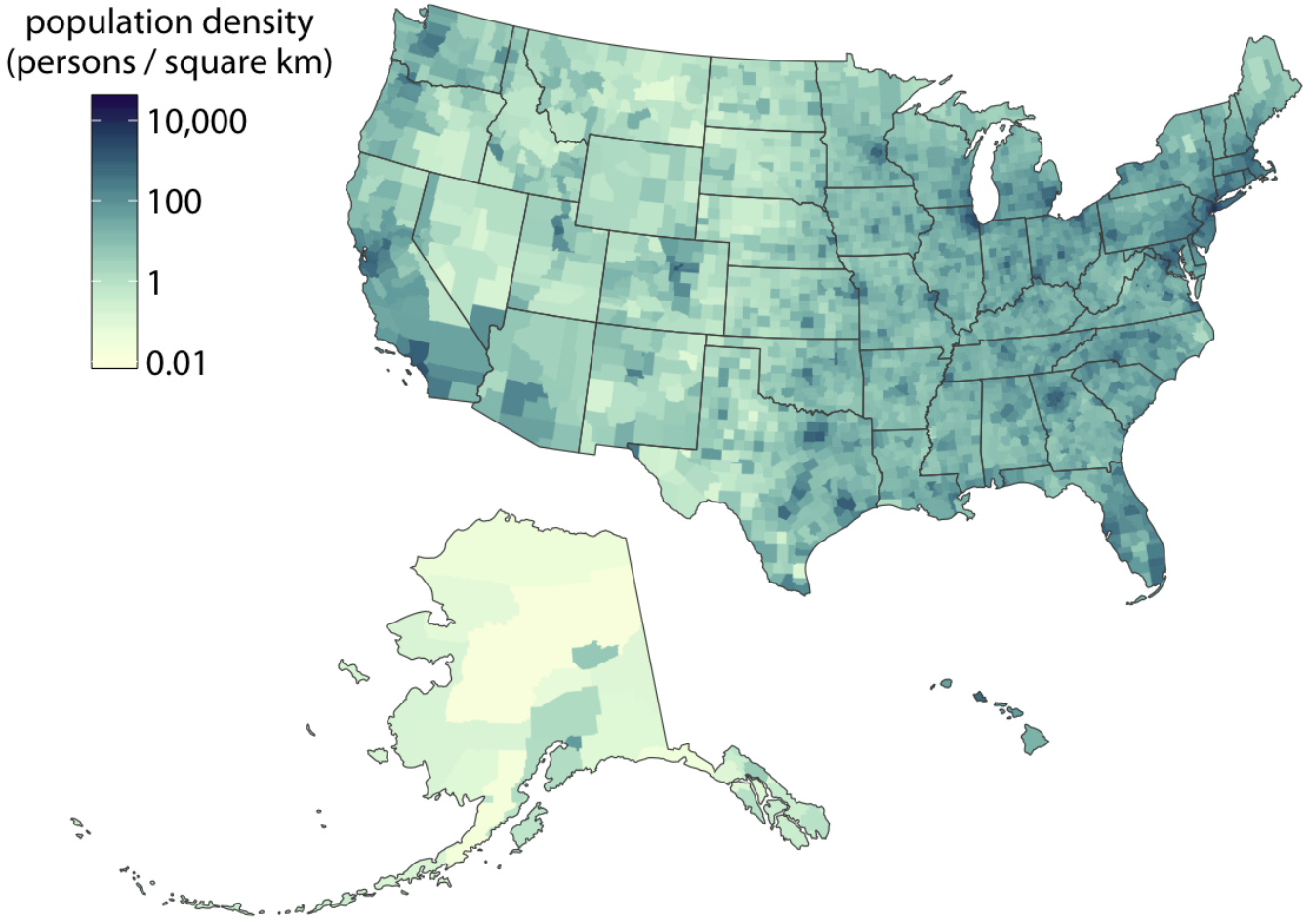


city labels, scale bar

Mapping to aesthetics

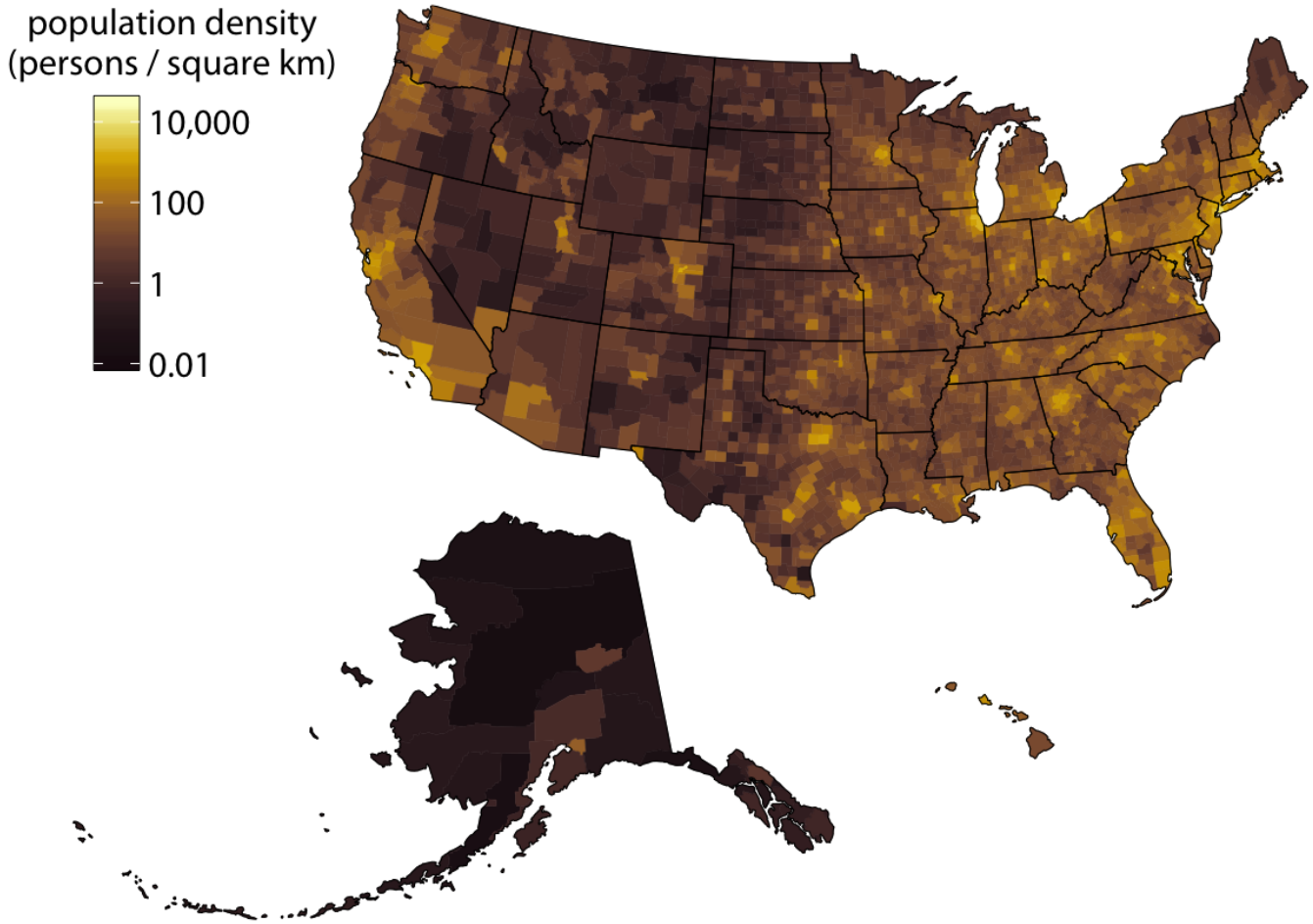


Choropleth mapping



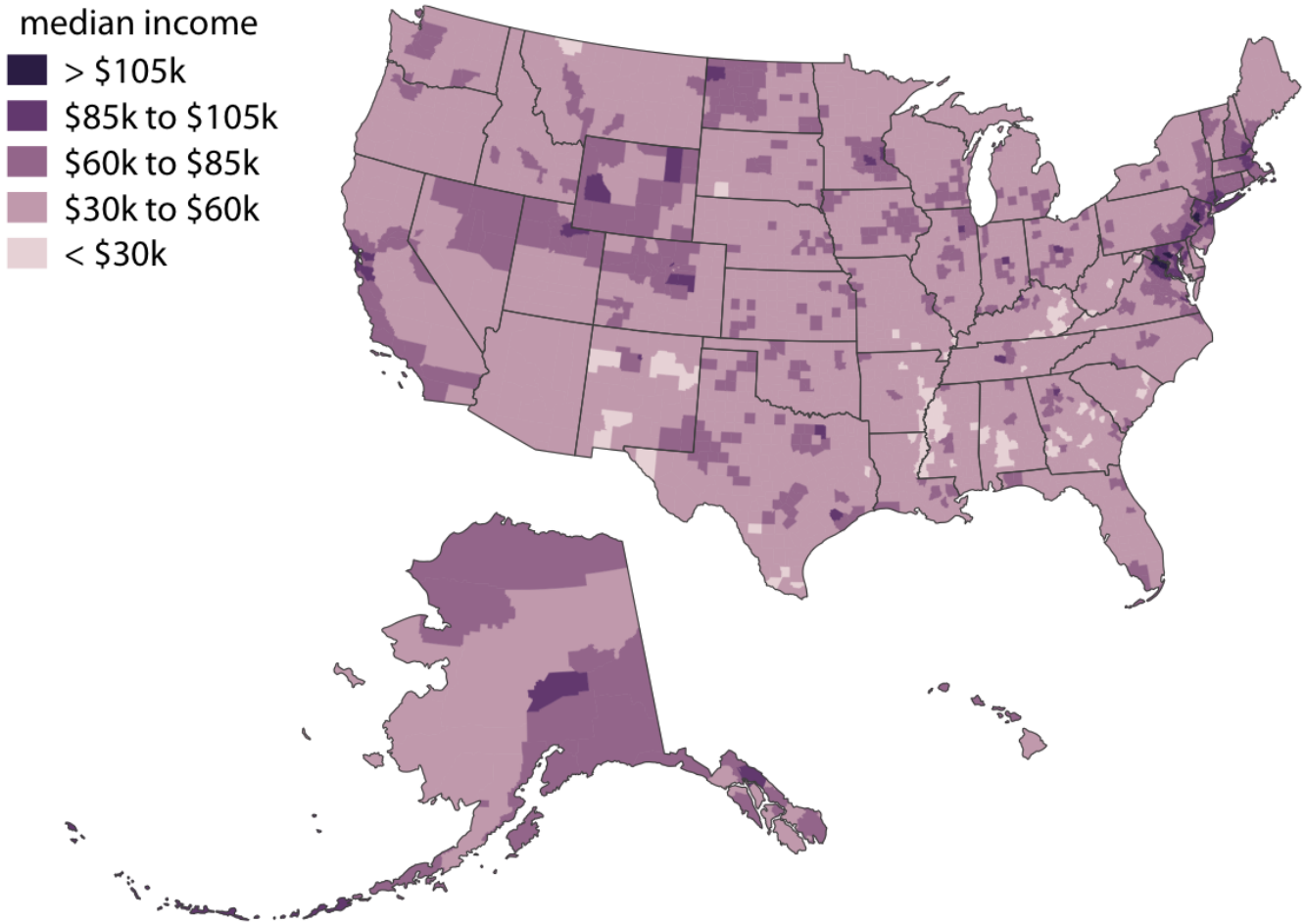
- Work best for **density**.
- Darker colors generally mean higher density.

Choropleth mapping



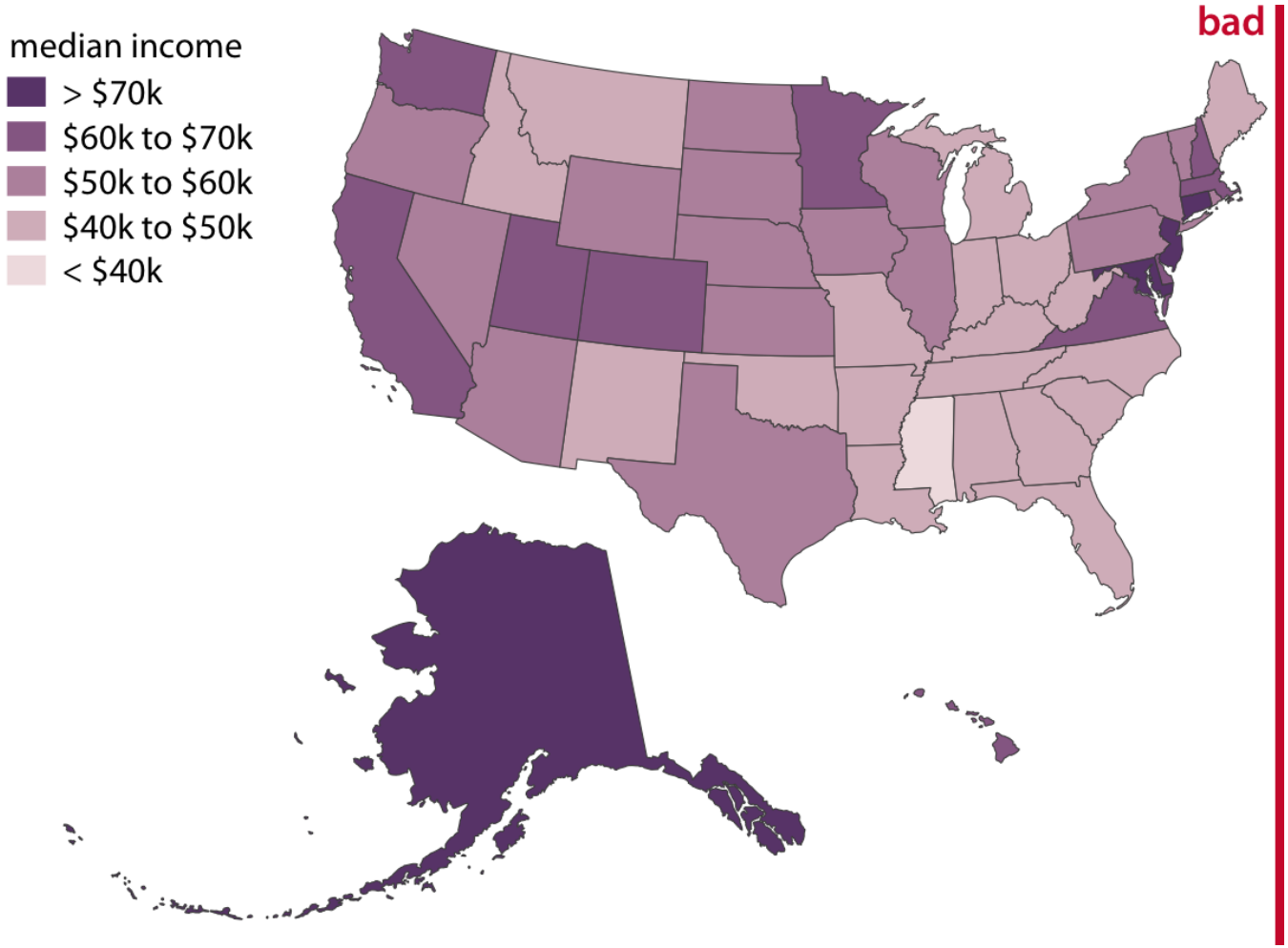
- Sometimes light colors "shine out" from the background.
- Keep colors in the red-yellow spectrum for this purpose.
- Generally, darker colors look better on white paper.

Discrete colors



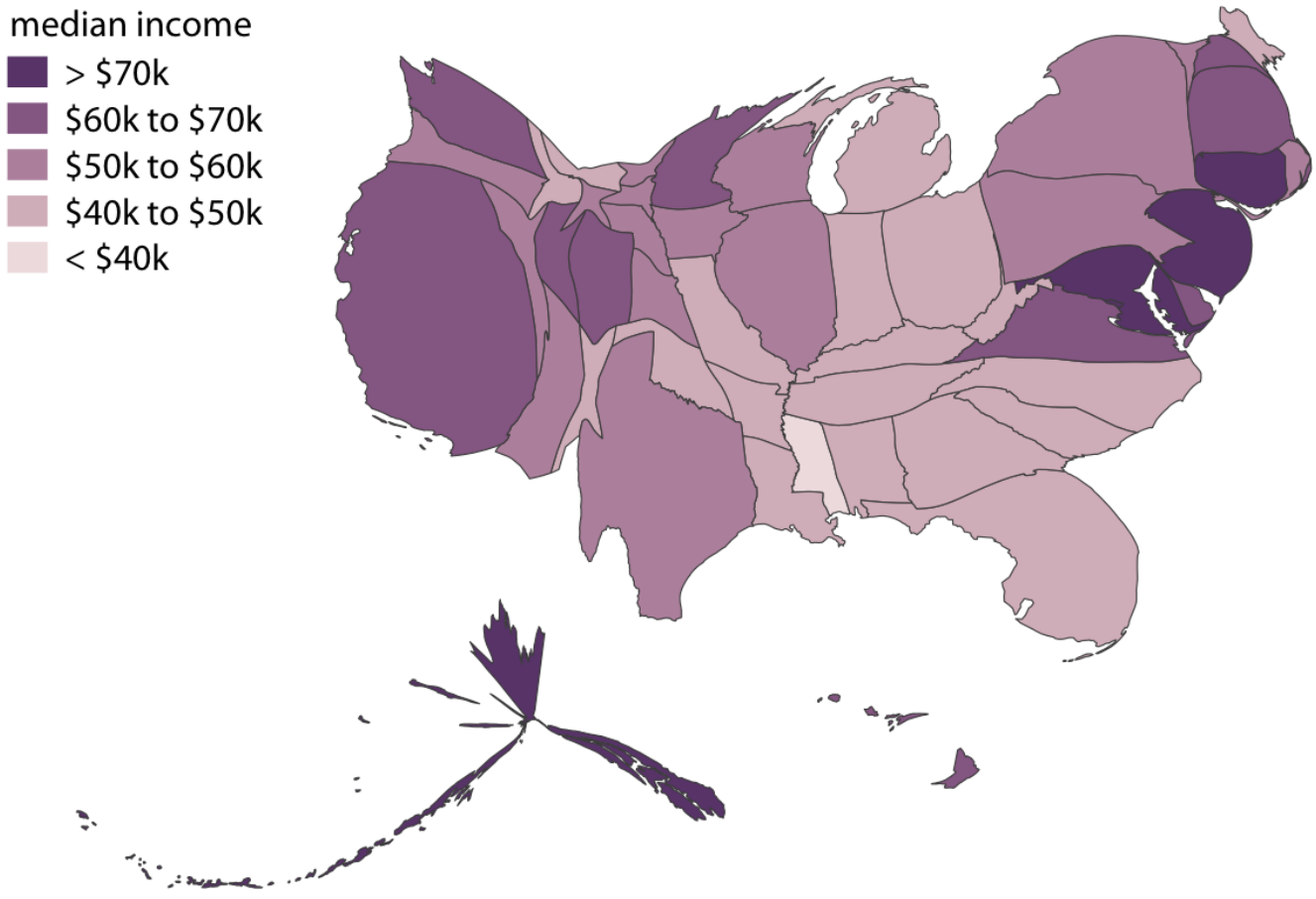
- Humans are bad at specific colors.
- Binning the data makes colors contrast more.
- Counties are approximately equal in size.

Areas not representational



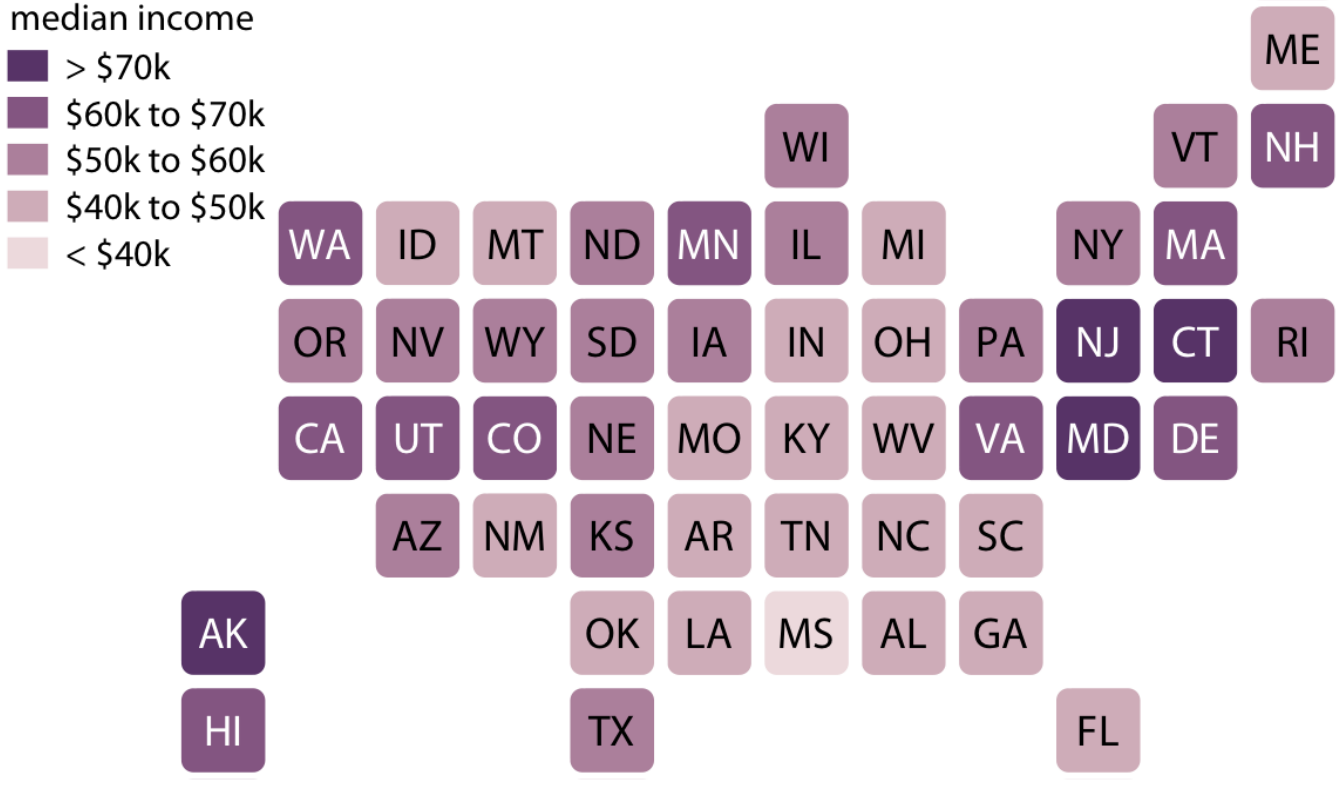
- Alaska dominates, but with very sparse population.

Cartograms



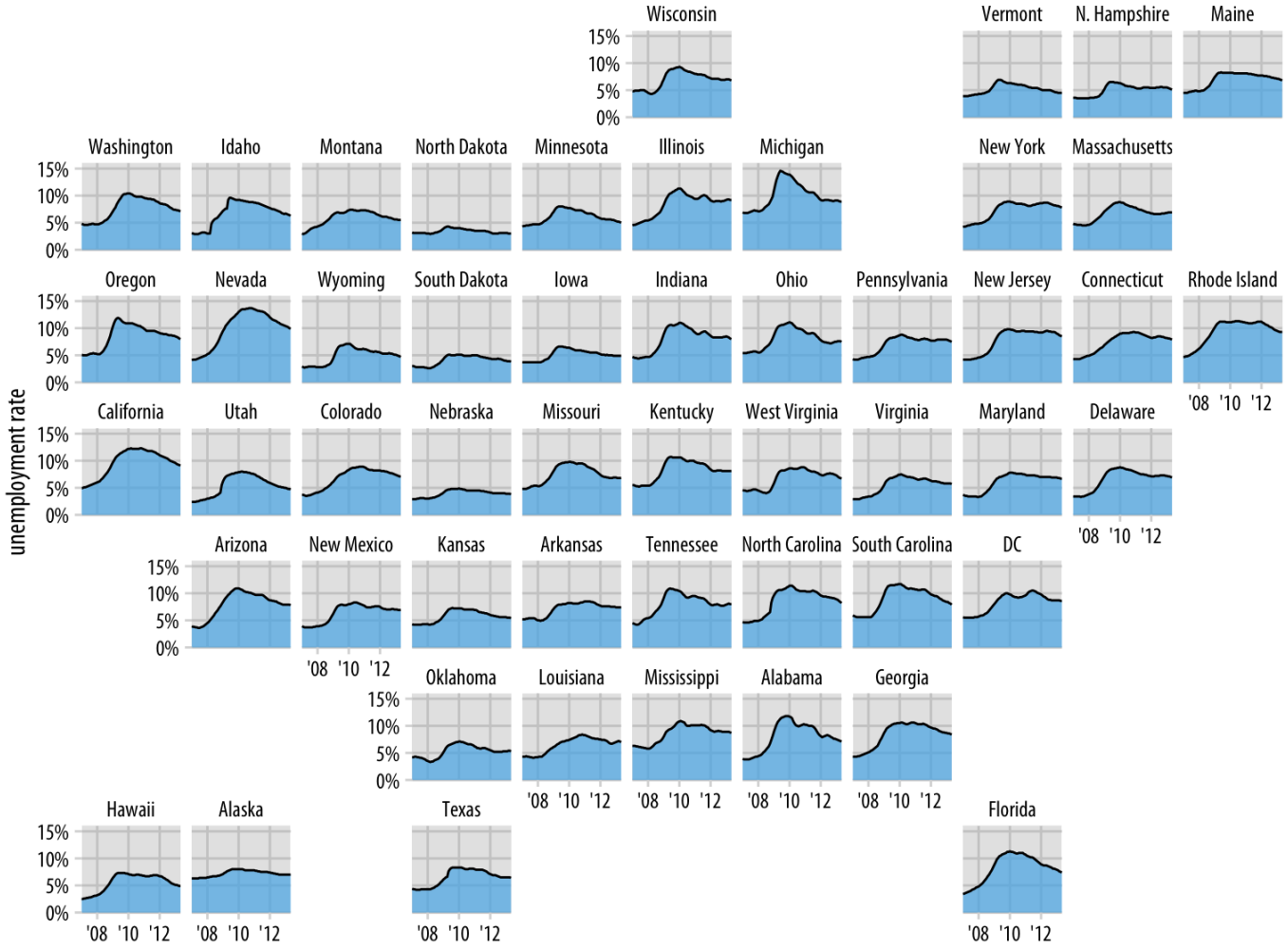
- Change areas to be proportional to population.

Cartogram heatmap



- Areas not proportional, but states don't stand out based on area.

Cartogram heatmap



- Works better than alphabetical order.
- Easier to find states.
- States in similar areas similar.