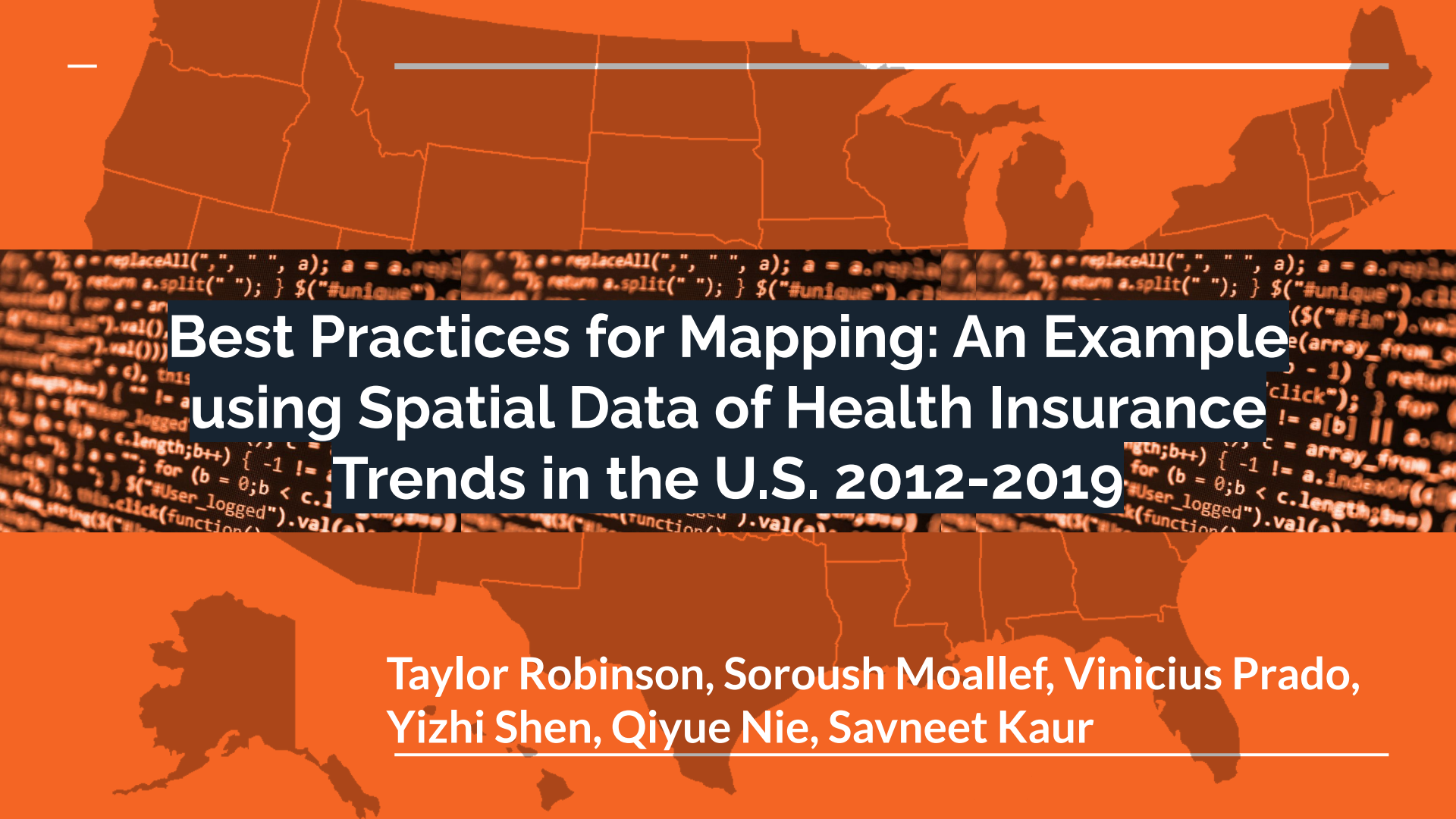
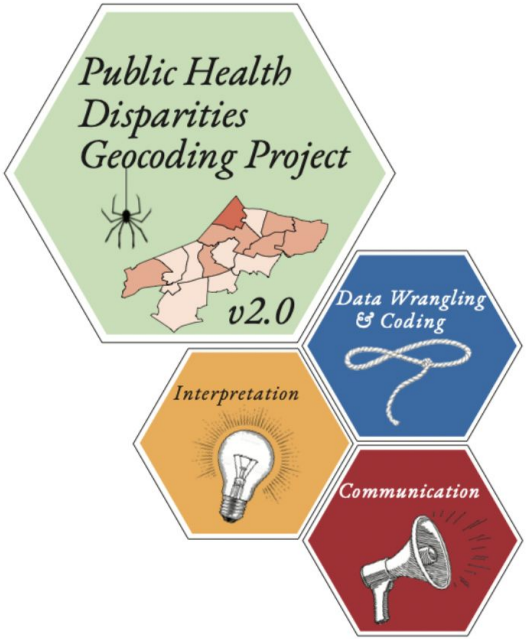


Best Practices for Mapping: An Example using Spatial Data of Health Insurance Trends in the U.S. 2012-2019

Taylor Robinson, Soroush Moallef, Vinicius Prado, Yizhi Shen, Qiyue Nie, Savneet Kaur



HARVARD T.H. CHAN | SCHOOL OF PUBLIC HEALTH
Powerful ideas for a healthier world

Outline

- Mapping Importance
- Data Considerations
- Best practices with data
- Spatial Analysis using Public Health Geocoding Project
- Best practices with mapping
- Communication Challenges

A+++ worthy research

Why does mapping matter?

- **Logistics, planning and resource allocation** visualization of the spatial distribution of healthcare resources and patient density
- **Emergency response and disaster management** by providing real-time information
- **Monitor changes** in **healthcare** services and infrastructure - **access and utilization**
- **Understand and address** healthcare **disparities and inequities**, in terms of access to healthcare services

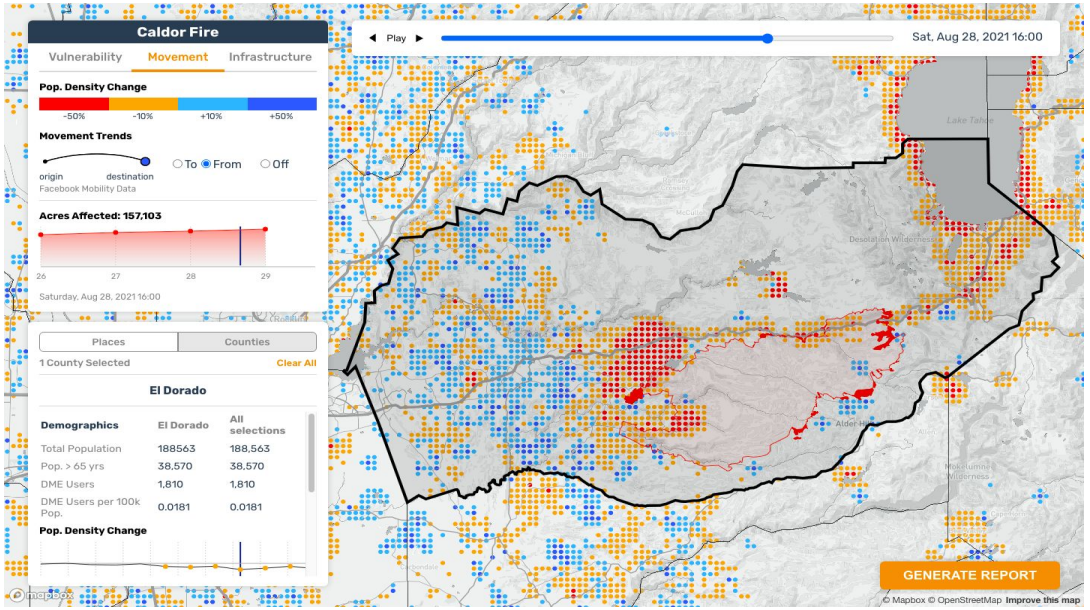


Fig. Changes in population density during the Caldor Fire in El Dorado, CA, Aug 2021, inferred from aggregated mobile device utilization data*

*Andrew Schroeder et al. CrisisReady's novel framework for transdisciplinary translation: Case-studies in wildfire and hurricane response

Data Considerations

Data Considerations for best practices

- Three main types of data:
  - Points
  - Lines
  - Polygons
- Census data can be provided in the form of boundaries (polygones) or centroids (point).
-

Geographic Entities

Legal/Administrative

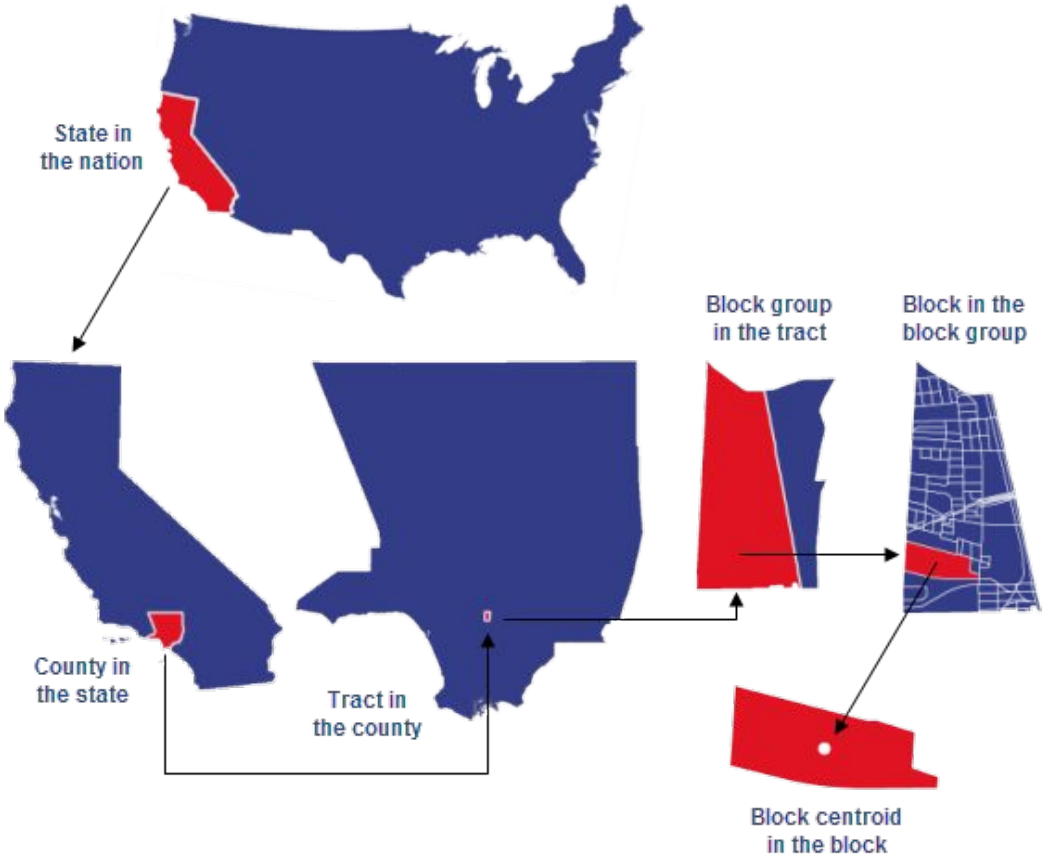
- States
- Counties
- Minor civil divisions
- Congressional districts
- School districts
- Incorporated places

Statistical

- Census tracts
- Metropolitan/Micropolitan statistical areas
- Urban areas
- Census designated places

Which statistical geographic unit are you interested in?

From:
https://learn.arcgis.com/en/related-concepts/united-states-census-geography.htm



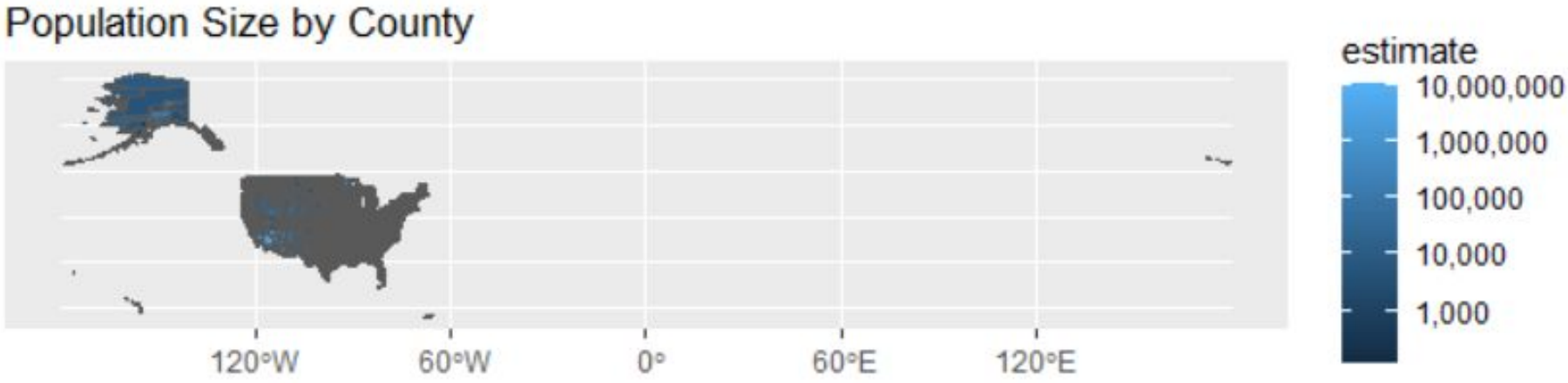
Other considerations when preparing our spatial data from ACS

- Data dictionary
  - Good availability: ACS PUMS Data Dictionary
- Data cleaning
  - Geocoding with Google Maps API
  - Removing some messy descriptions
  - Recoding some variables
  - Dealing with missing data
- Specifying longitudinal and latitudinal coordinates

Spatial Visualization: Example

Displaying the data

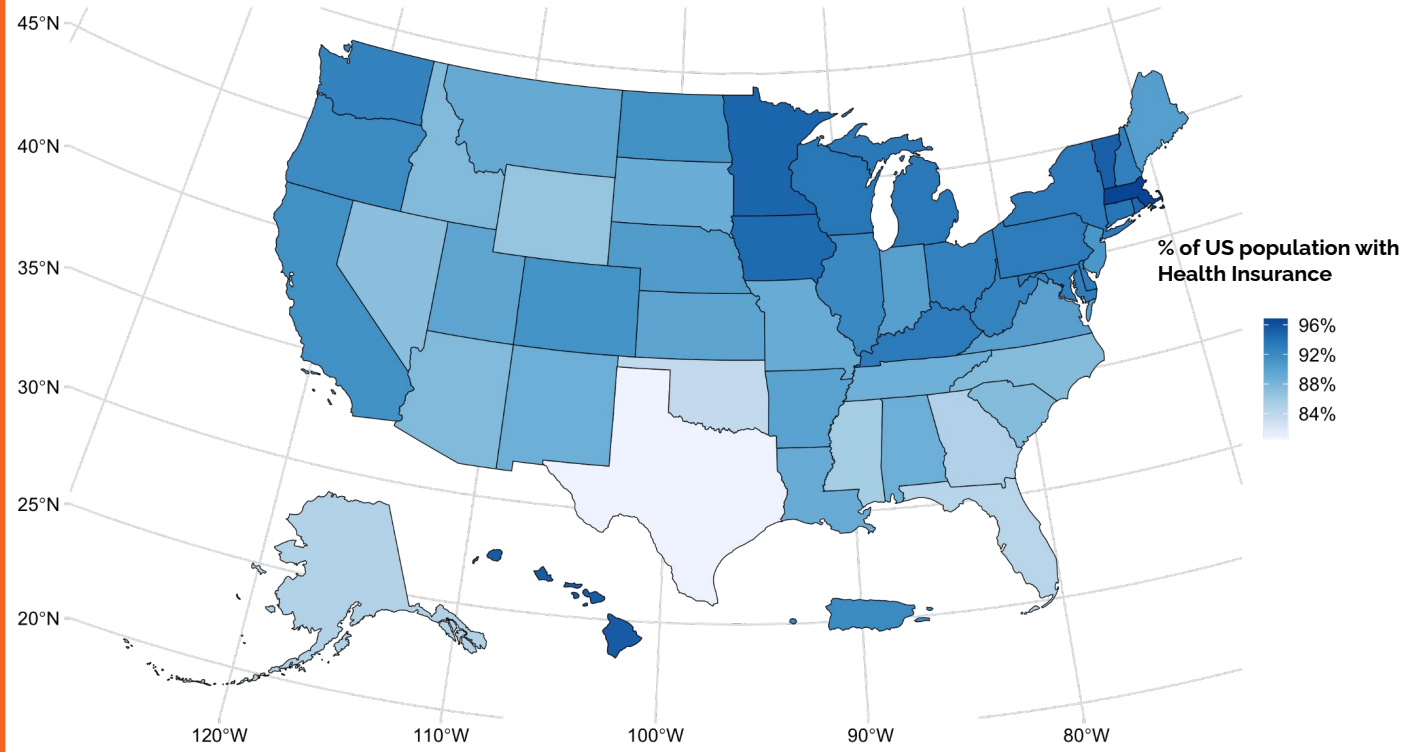
Is this a good map?



Inappropriate scale: `tigris::shift_geometry()`

Inappropriate color theme: `scale_fill_distiller(direction = 1)`

Percent U.S. Population with Health Insurance (2019)



Data Source: American Community Survey (ACS), U.S. Census Bureau (2012-2019)

Public Health Disparities Geocoding Project v2.0

Data Wrangling & Coding

Interpretation

Communication

https://phdgp.github.io/PHDGP2.0

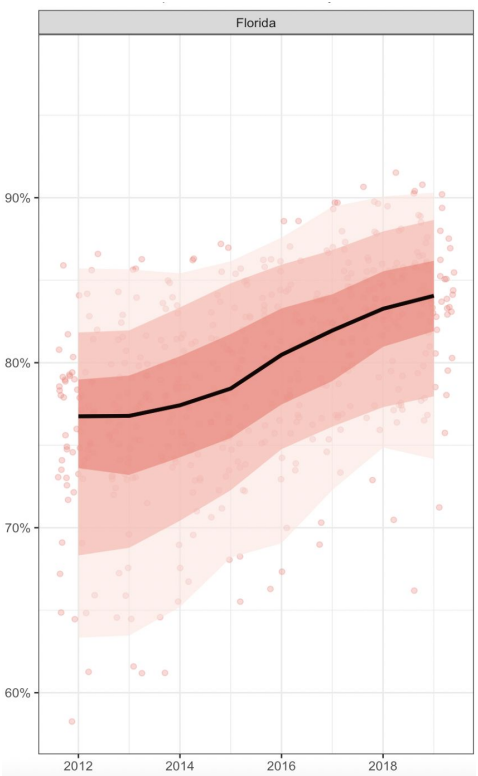
HARVARD T.H. CHAN SCHOOL OF PUBLIC HEALTH
Powerful ideas for a healthier world

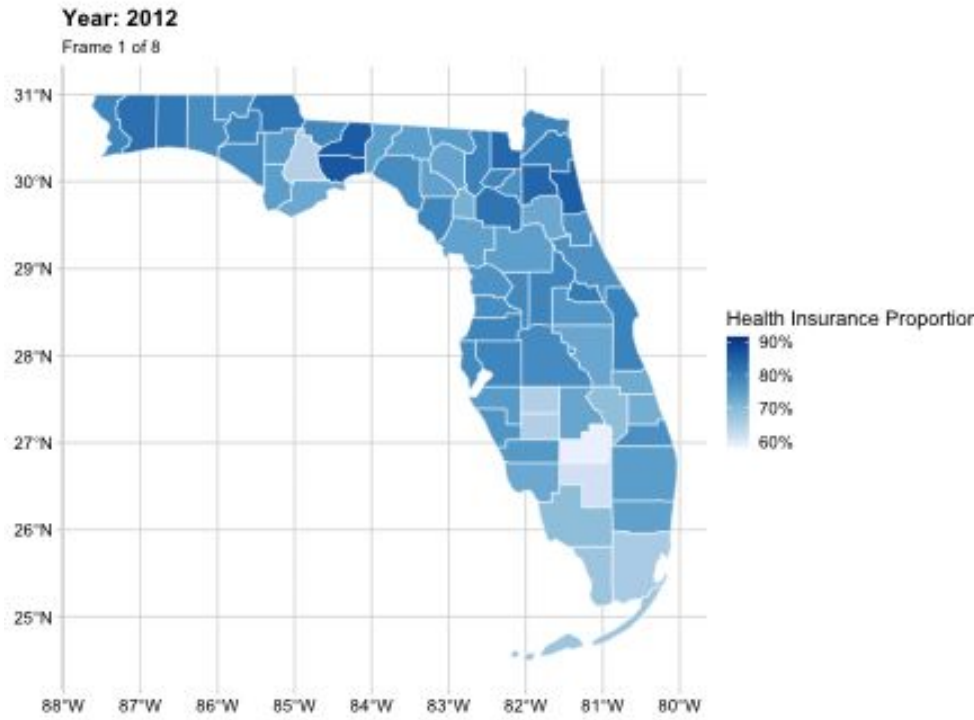
Testa C, Chen JT, Hall E, Javadi D, Morgan J, Rushovich T, Saha S, Waterman PD, Krieger N. The Public Health Disparities Geocoding Project 2.0. Training Manual. Available as of October 30, 2022

Spatial Analysis

Is there within state difference in health insurance?

Florida Health Insurance Proportion (2012-2019)



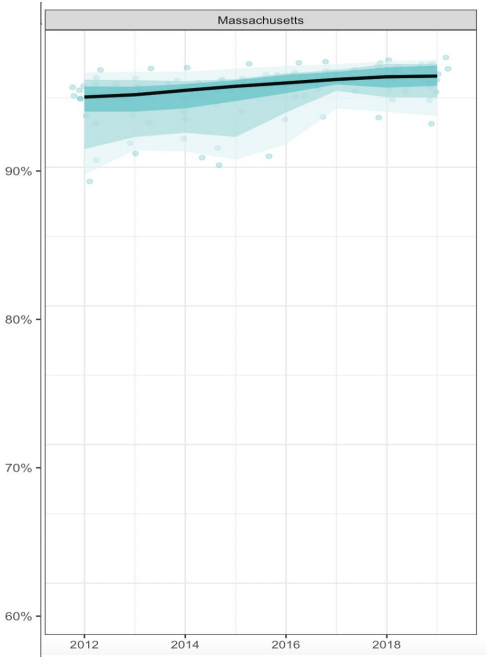


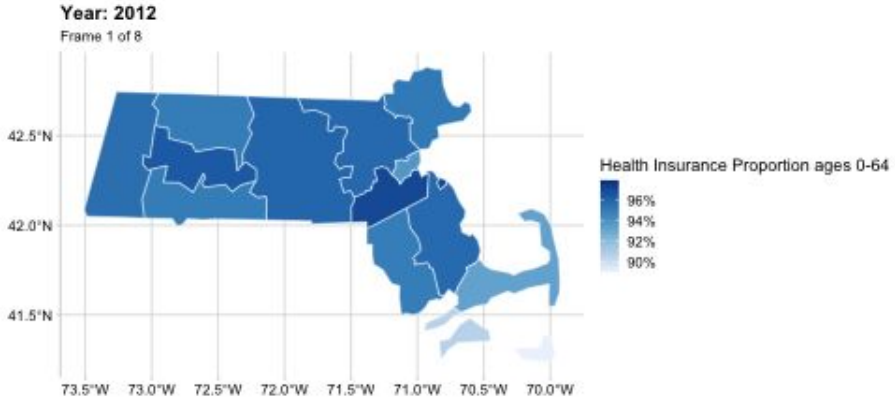
Data Source: American Community Survey (ACS), U.S. Census Bureau (2012-2019)

Massachusetts Health Insurance Proportion (2012-2019)

gganimate package

```
b +
transition_time(year)
```

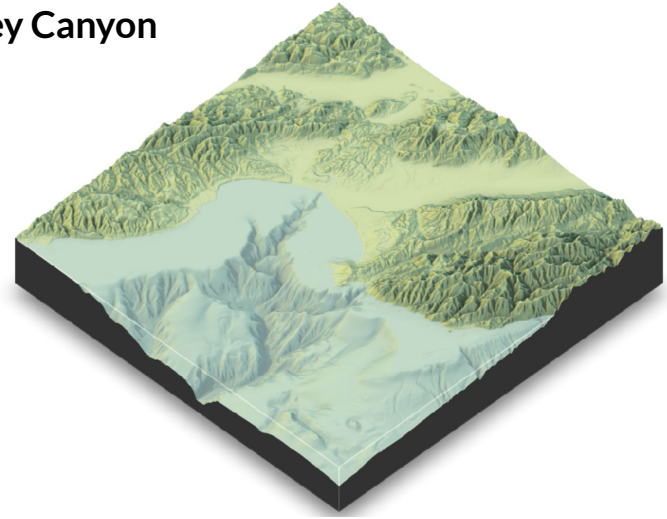




Data Source: American Community Survey (ACS), U.S. Census Bureau (2012-2019)

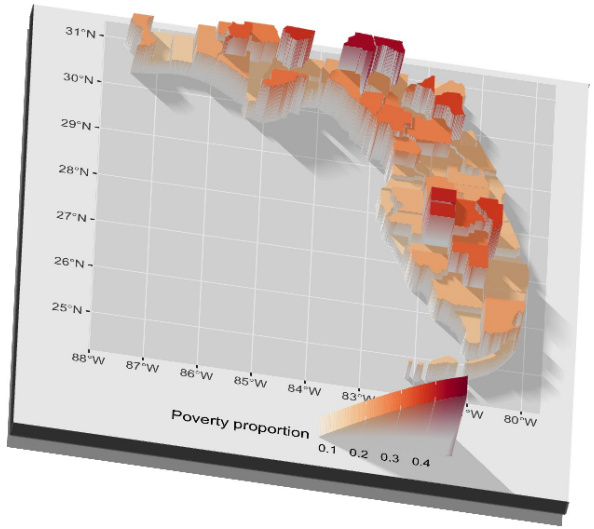
3D Maps - quick comments

Monterey Canyon System



Source: rayshader R package
Available at:www.rayshader.com

Florida Poverty Proportion - 2019



Data Source: American Community Survey (ACS), U.S. Census Bureau (2019)

Data Practices for Mapping Summary

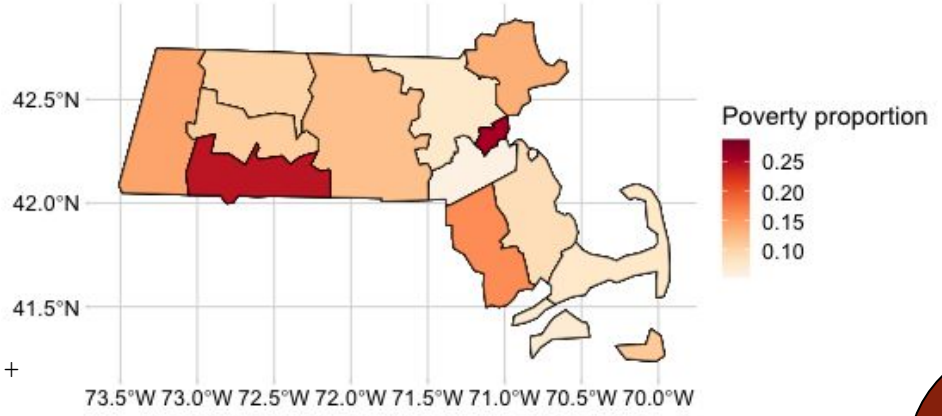
Thorough data dictionaries help to prevent confusion

Theory & theoretical frameworks are important to structure thinking about spatial relationships

It's important to check on missing data and the populations represented

It's Important to Specify the Correct Projection

```
ggplot(health_insurance_sf_ma) +
  geom_sf(
    aes(fill = poverty_prop),
    color = "black", size = 0.1
  ) +
  scale_fill_continuous_sequential(
    palette = "OrRd", rev = TRUE
  ) +
  coord_sf() +
  labs(fill = 'Poverty proportion') +
  theme_minimal_grid(11)
ggsave("proj01_stand.png")
```



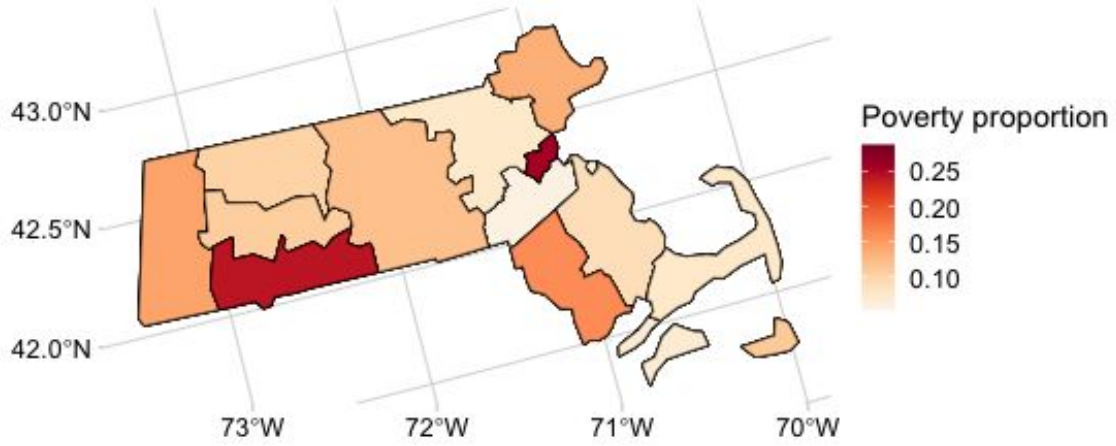
Goal: Choose the least distorted projection

Your projection displays the coordinate system (it is a transformation of latitude and longitude). Projection selection can distort the area and calculations in spatial analysis.

coord_sf() helps to select a common coordinate reference system (CRS)

It's Important to Specify the Correct Projection

```
ggplot(health_insurance_sf_ma) +
  geom_sf(
    aes(fill = poverty_prop),
    color = "black", size = 0.1
  ) +
  scale_fill_continuous_sequential(
    palette = "OrRd", rev = TRUE
  ) +
    coord_sf(
    # Texas Centric Albers Equal Area
    crs = 3083
  ) +
  labs(fill = 'Poverty proportion') +
  theme_minimal_grid(11)
ggsave("proj01_stand.png")
```



Mapping is a Communications Challenge

Be clear about the purpose of your map and the intended audience

Use dynamic color schemes and features to highlight key points

Maps are subjective. Make sure associations seen show the real story

Conclusions

Why does best practice in mapping matter and how we have grown

- Importance of best practice
  - Showcase expertise of map-maker.
  - Deliver the exact spatial message to the audience
    - You cannot ask “why” any more
    - You cannot hear “want” in the map message
- Our growth
  - We make our maps recognizable and easy to read with:
    - Purposeful color theme
    - Carefully chosen projection system
  - We make our maps informative and effective with:
    - Clarity of all labels
    - GIF Animation
    - Motivated comparison
    - 3D maps

Link to our mapping deliverable

https://github.com/viniciusdopradomonteiro/Mapping_Practice

Follow us on GitHub
@viniciusdopradomonteiro
@taylorjrobinson
@smoallef

Resources to learn more

1. Chapter 9 Making Maps with R, Geocomputation with R:

   https://geocompr.robinlovelace.net/adv-map.html#adv-map

   Main Takeaway:

   Static map -> animated map -> interactive map

   Related package: tmap, geom_sf() with ggplot2; gganimate (which builds on ggplot2); mapview

2. R tutorial: Creating Maps and mapping data with ggplot2: https://www.youtube.com/watch?v=AgWgPSZ7Gp0&ab_channel=StatisticsGuideswithDrPaulChristiansen

3. Chapter 8 Plotting Spatial Data, Spatial Data Science:

   https://r-spatial.org/book/08-Plotting.html

Other mapping resources

- Columbia University
  - Spatial Epidemiology Summer Course (Epi summer)
- University of Washington Summer Institute Summer Institute in Statistics and Modeling in Infectious Diseases
  - Spatial Statistics w/ Lance Waller and Jon Wakefield
  - Novel Data Streams
- Applied Spatial Statistics Book by Lance Waller
- Twitter !
  - Hari Iyer - Cancer GIS Workshop