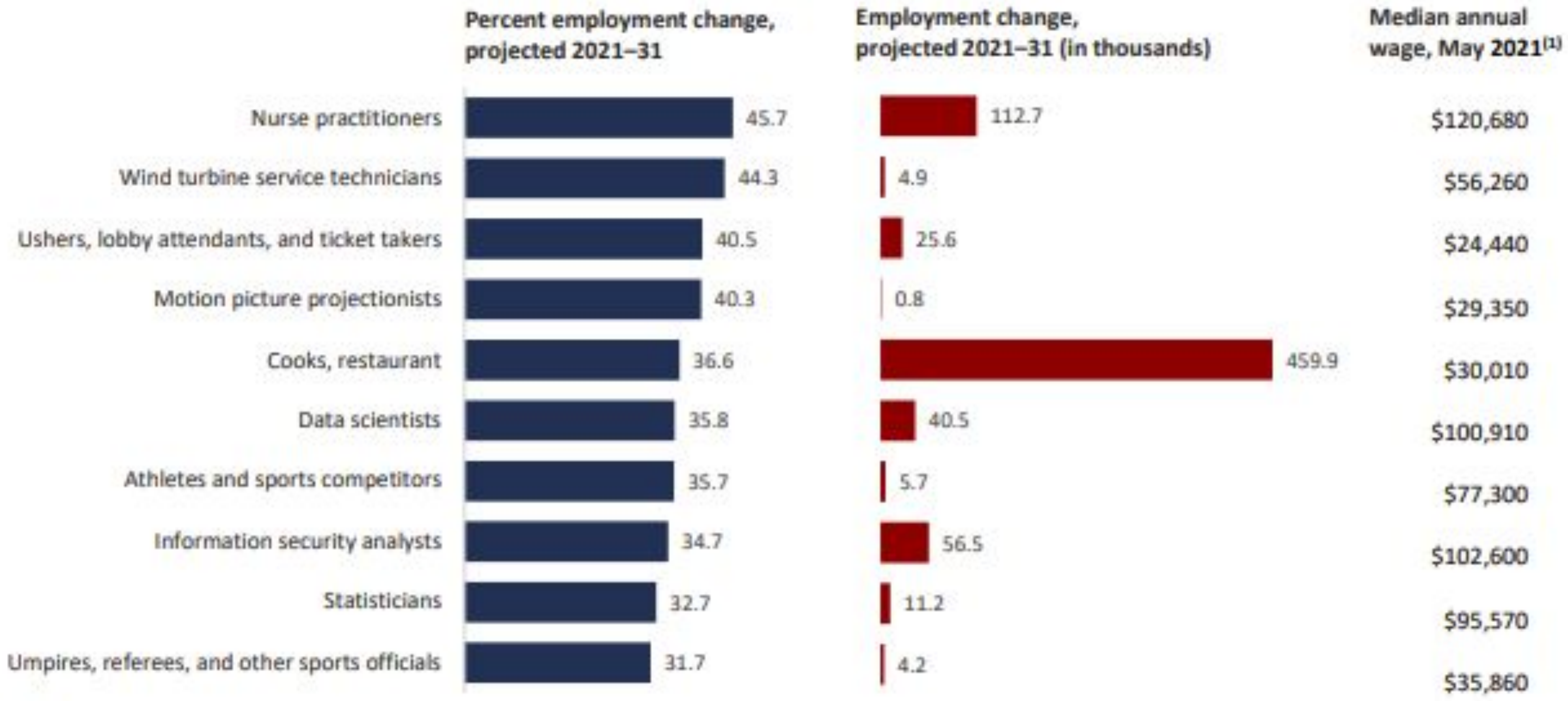
Chart 6. Ten fastest growing occupations, projected 2021–31



Footnotes:

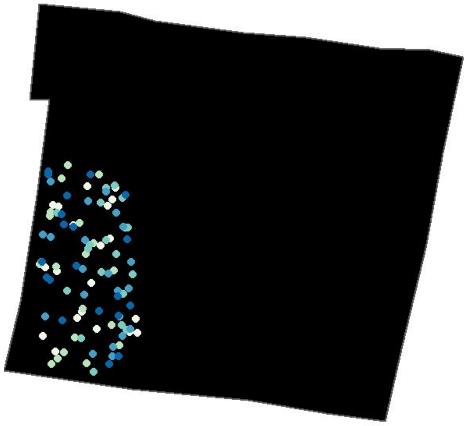
1) Data are from the Occupational Employment and Wage Statistics (OEWS) program, U.S. Bureau of Labor Statistics. Wage data cover non-farm wage and salary workers and do not cover the self-employed, owners and partners in unincorporated firms, or household workers.

Note: Some occupations are expected to experience cyclical recovery from the COVID-19 recession, which results in fast projected growth for these occupations.

Source: U.S. Bureau of Labor Statistics

On the importance of indicators - Transit Oriented Development

Public Policy Analytics - Week 2



Agenda:

1. Syllabus and course logistics
    a. Github
2. Notes/questions regarding homework
3. Class 1 review
4. Let’s talk Indicators
5. Sources of bias in spatial indicators
6. Geoprocessing in R
7. Assignment - TOD

Big Theme 1:

Algorithmic decision-making is just a new take on the traditional approach governments use to design programs

Big Theme 2:

Evidence-based, proactive decision-making is preferable to reactive decision-making

Big Theme 3:

Critical thinking, transparency, and communication are key elements to the analytical process.

Context > Engineering chops

Big Theme 4:

The importance of spatial thinking

What does the *spatial process* refer to?

Indicators

Average revenue per passenger ride

U.S. 2015; Sorted by total system trips. Includes heavy rail (subway) and light rail.

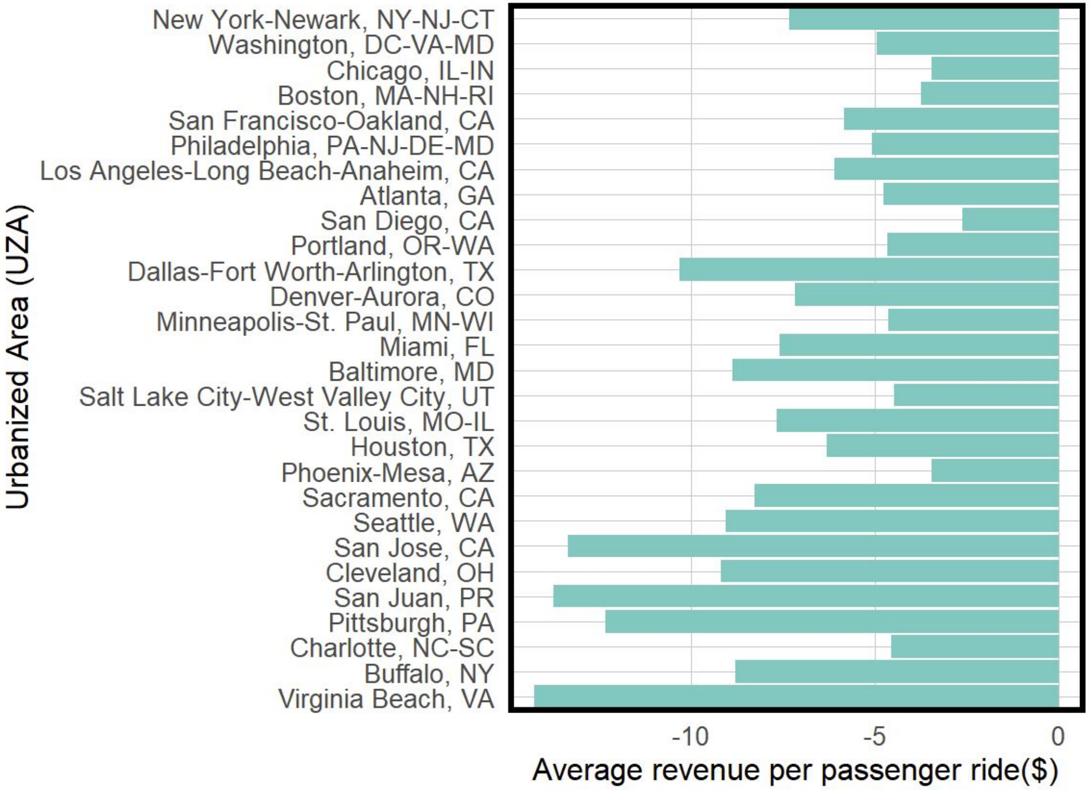


Figure 1.1

What makes a good indicator?

A good indicator is **relatable** - motivated by a pressing policy concern.

“How is it possible that passenger rail in and around New York City has such widespread delays, service suspensions and rider discontent?”

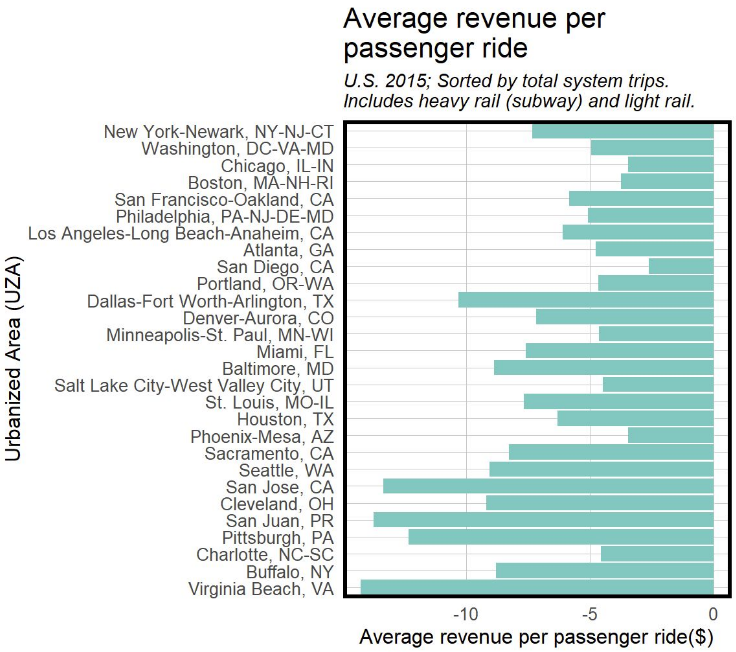


Figure 1.1

What makes a good indicator?

A good indicator is **simple** -

Means, sums and counts. Simplicity helps the audience understand the indicator's significance and keeps them engaged in the analysis.

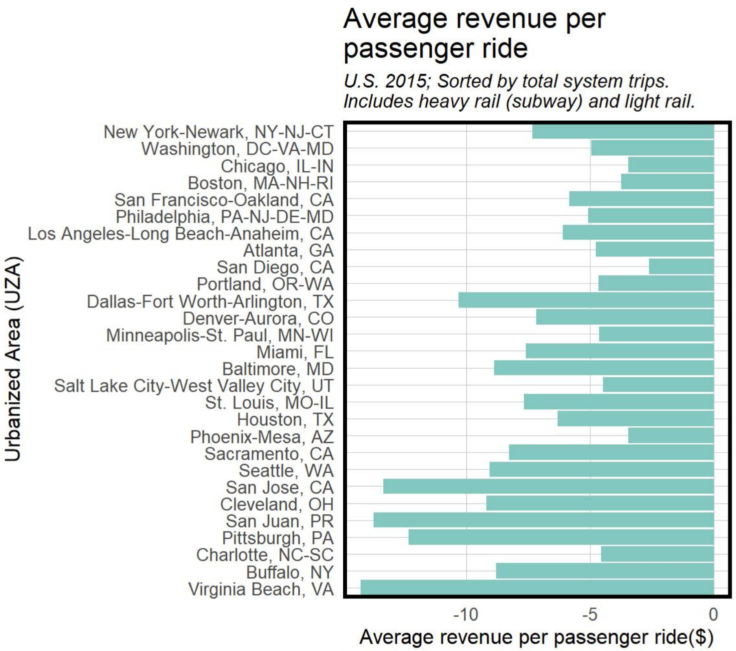


Figure 1.1

What makes a good indicator?

A good indicator is **relative** - and draws contrasts.

"How can New York City passenger rail, with the most trips, still lose more money (per passenger) than each of the next ten largest cities?"

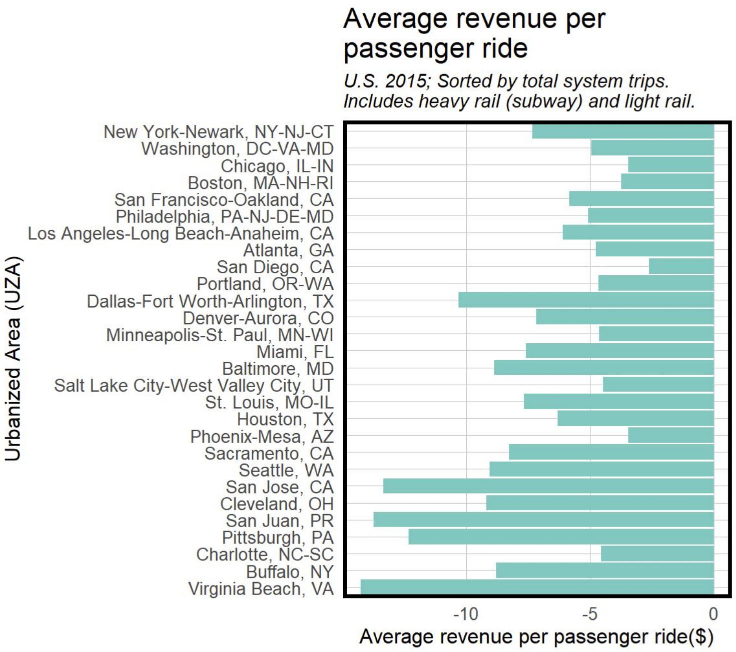
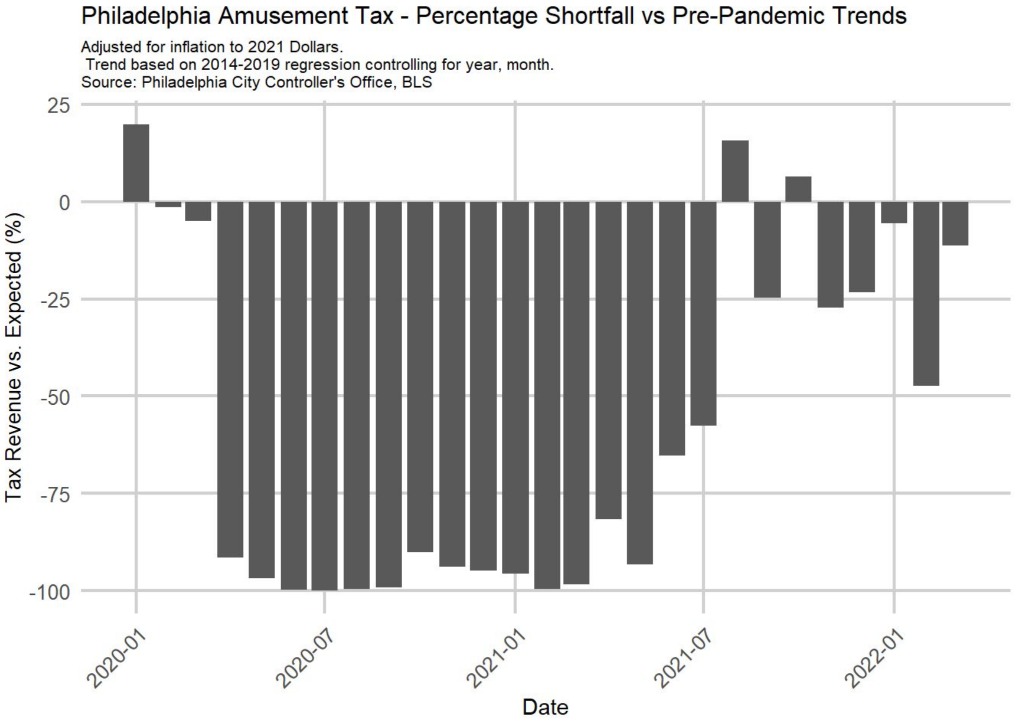


Figure 1.1

What makes a good indicator?

A good indicator is **relative** - and draws contrasts.



Fichman et al 2022 - https://mafichman.github.io/ACTF_nightlife/

What makes a good indicator?

A good indicator generates **more questions** than answers -

A good indicator fits into a broader *narrative* which helps motivate a more robust research agenda and ultimately, more applied analytics.

“What do we do about all this…!?!?”

“Glad you asked…”

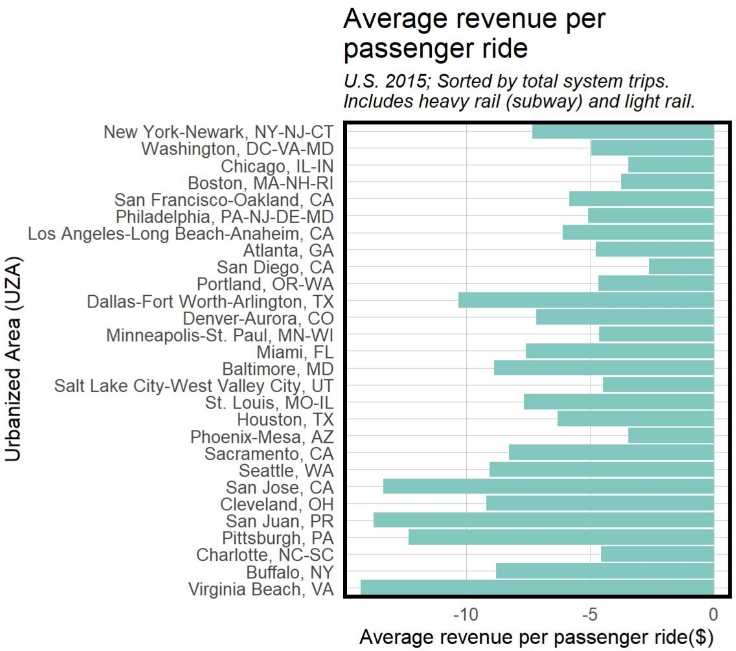


Figure 1.1

Exploring TOD with space/time Census indicators

What is TOD?

Average revenue per passenger ride

U.S. 2015; Sorted by total system trips. Includes heavy rail (subway) and light rail.

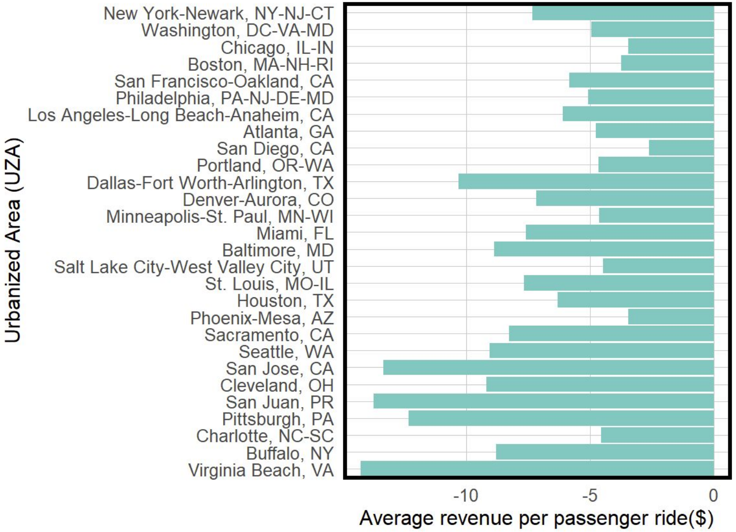


Figure 1.1

TOD

Our research question is -

Are renters willing to pay a rent premium to live close to transit?

Along the way, we will learn how assumptions can lead to **incorrect** policy conclusions.

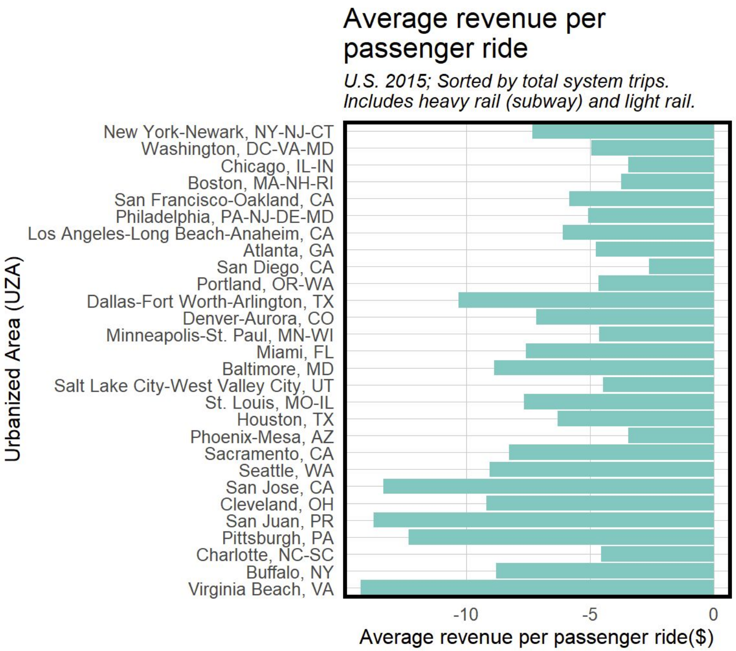


Figure 1.1

Mapping & Scale bias in areal aggregate data

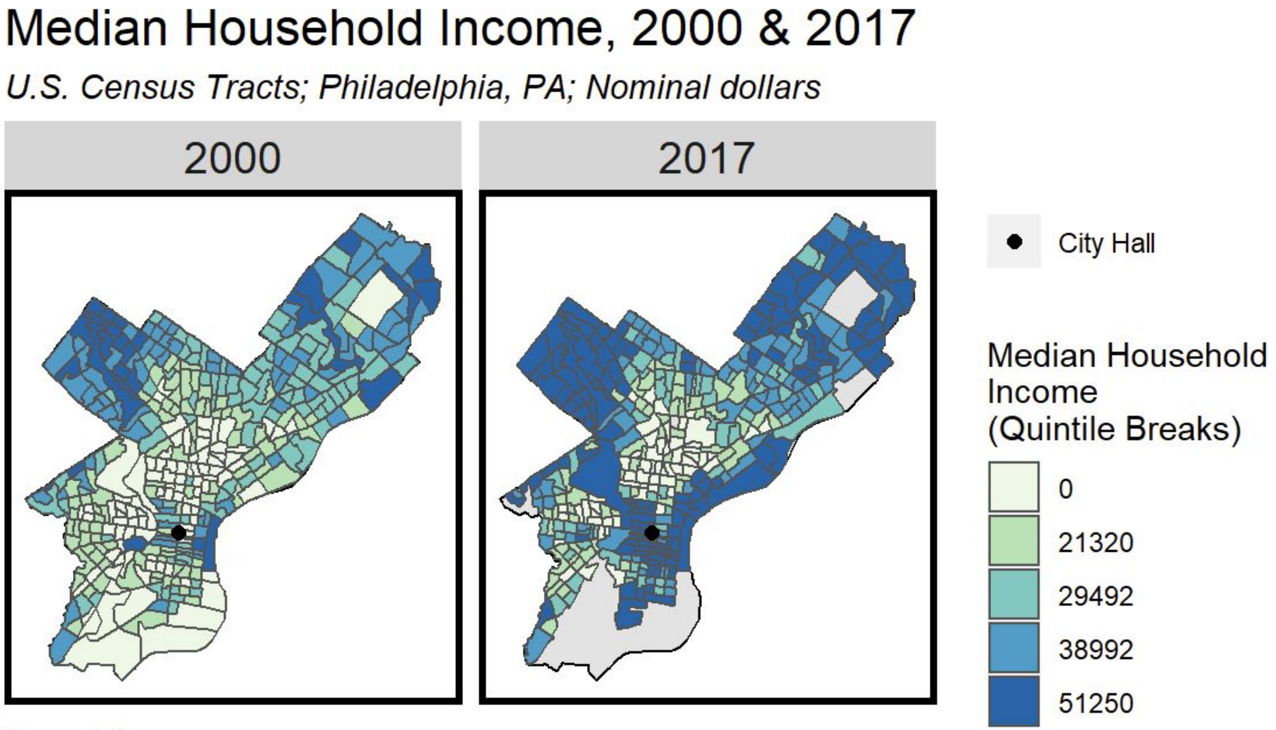
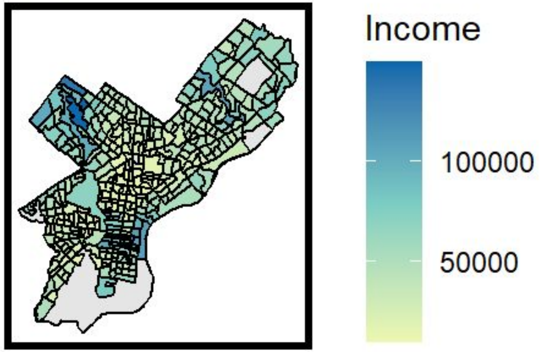
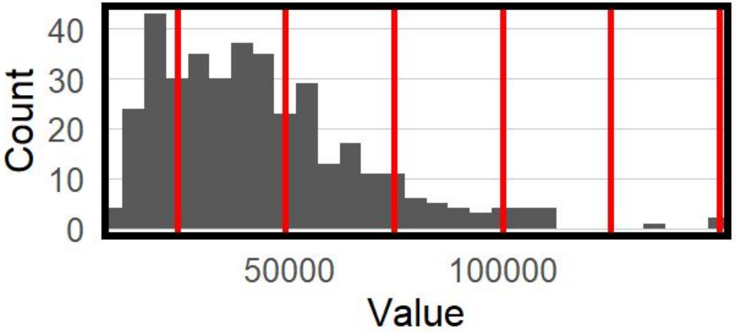


Figure 1.2

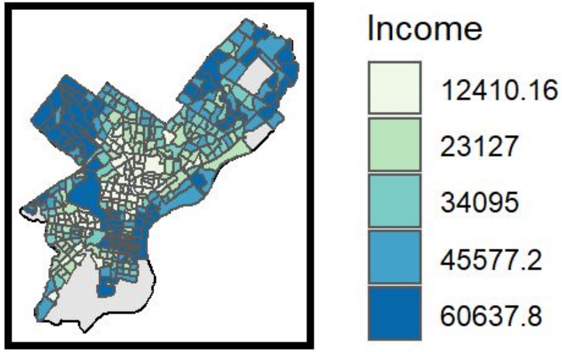
Map breaks

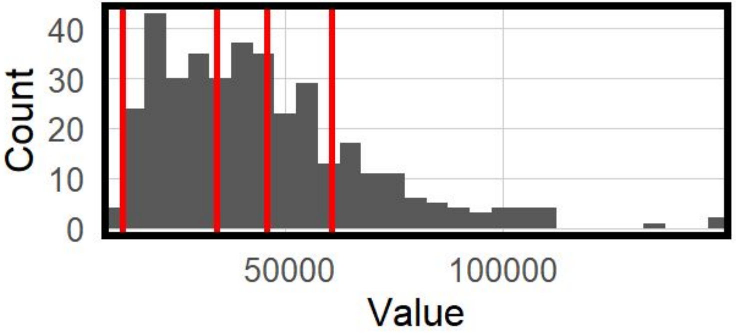
ggplot default breaks



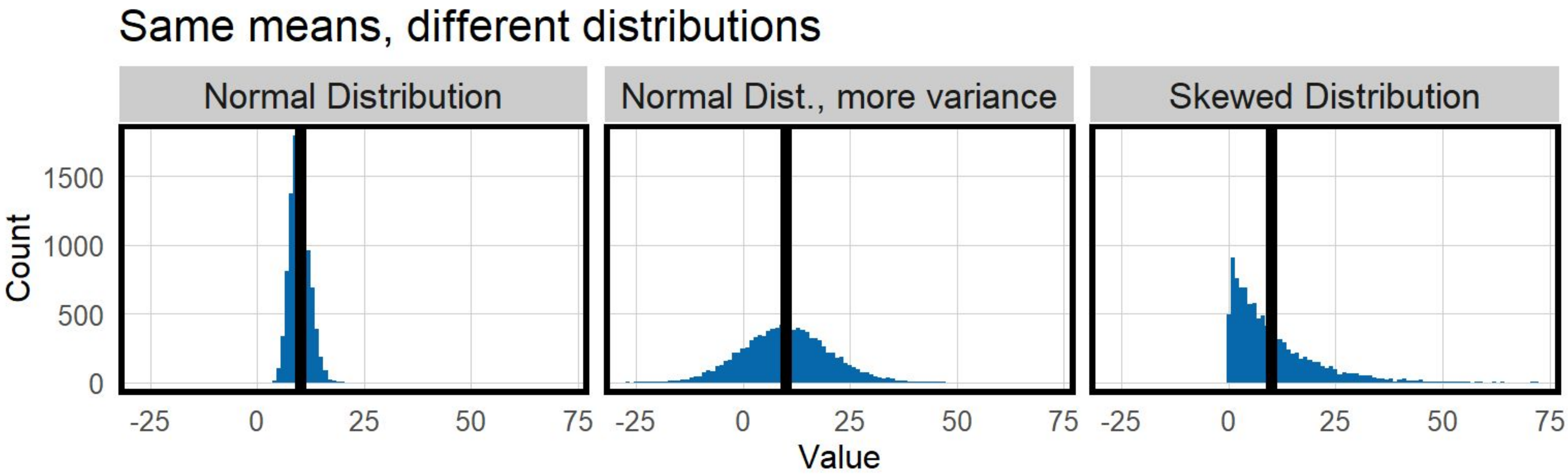


Quintile breaks

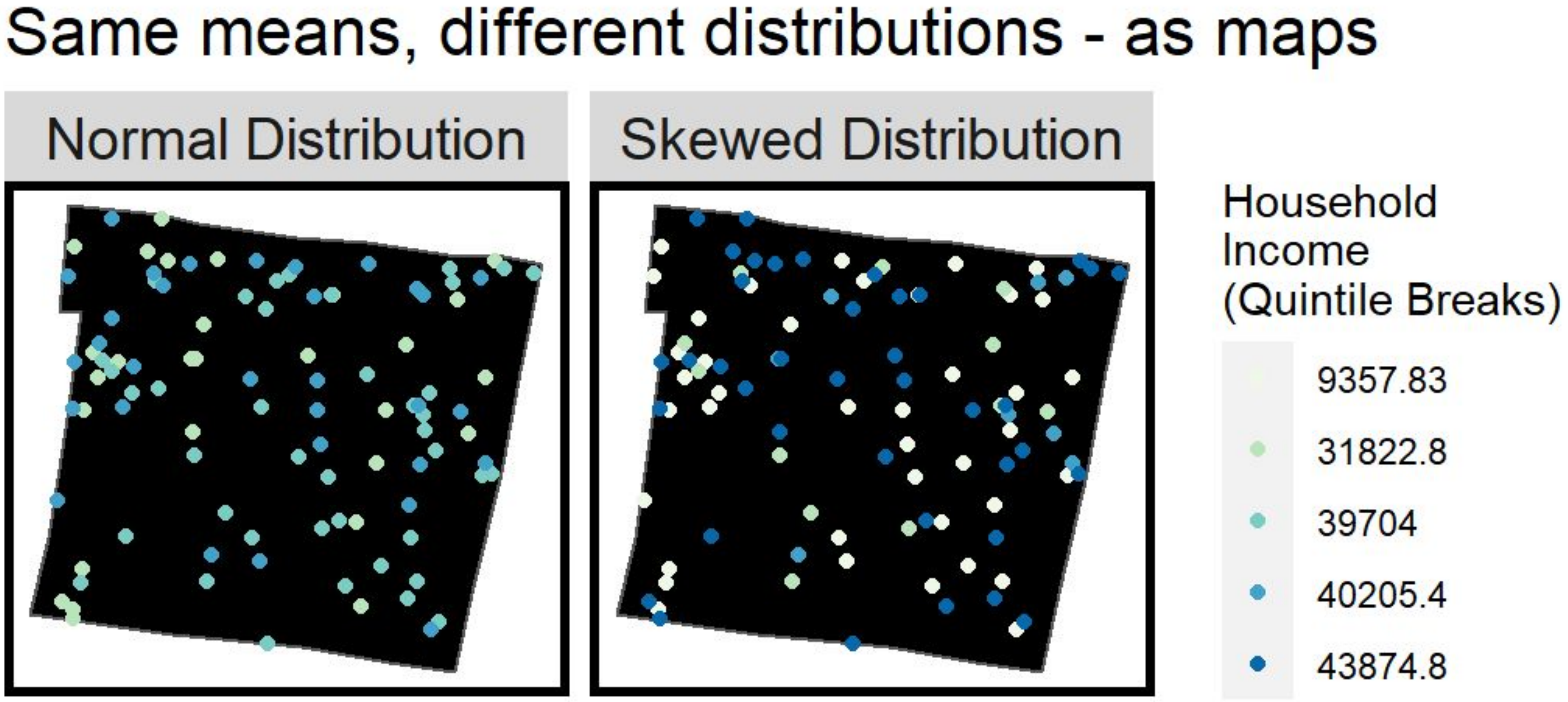




Scale bias - Ecological fallacy



Scale bias - Ecological fallacy



Scale bias - Modifiable Areal Unit Problem (MAUP)

Public Policy Analytics - Figure 2.5

Pay attention to scale bias…

Scale is the most difficult problem to overcome in spatial analysis

Bias in indicators

I know these points seem subtle and nerdy…

but…

Bias in indicators

1. I am nerd

2. More importantly, these biases turn into **assumptions**, that may ultimately affect your conclusions. You need to know how to interpret conclusions given these assumptions.

Important R packages in this chapter

Important R packages in this chapter

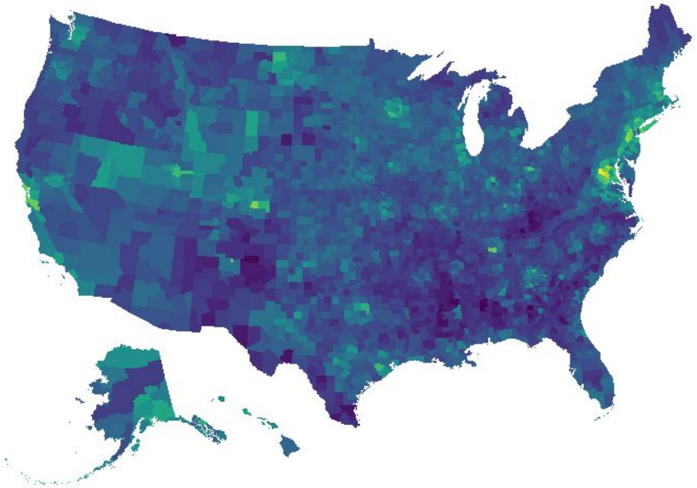
```
dinnersData %>%
    select(Name, Food)          %>%
    filter(Name == “Ken Steif”)
```

```
         Name                 Food
1 Ken Steif Chicken Parmesaen
```

Important R packages in this chapter

```
Simple feature collection with 1 feature and 2 fields
geometry type:  POINT
dimension:      XY
bbox:           xmin: -81.49826 ymin: 36.4314 xmax: -81.49826 ymax: 36.4314
geographic CRS: NAD27
        Name                  Food                    geometry
1 Ken Steif Chicken Parmesaen POINT (-81.49826 36.4314)
```

Important R packages in this chapter



Tidycensus

```
tracts00 <-
  get_decennial(geography = "tract", variables = c("P001001","P006002","PCT025050",
                                                    "PCT025009","P053001","H056001",
                                                    "P092001"),
                year=2000, state=42, county=101, geometry=T) %>%
  st_transform('ESRI:102728')
```

Make some maps in R

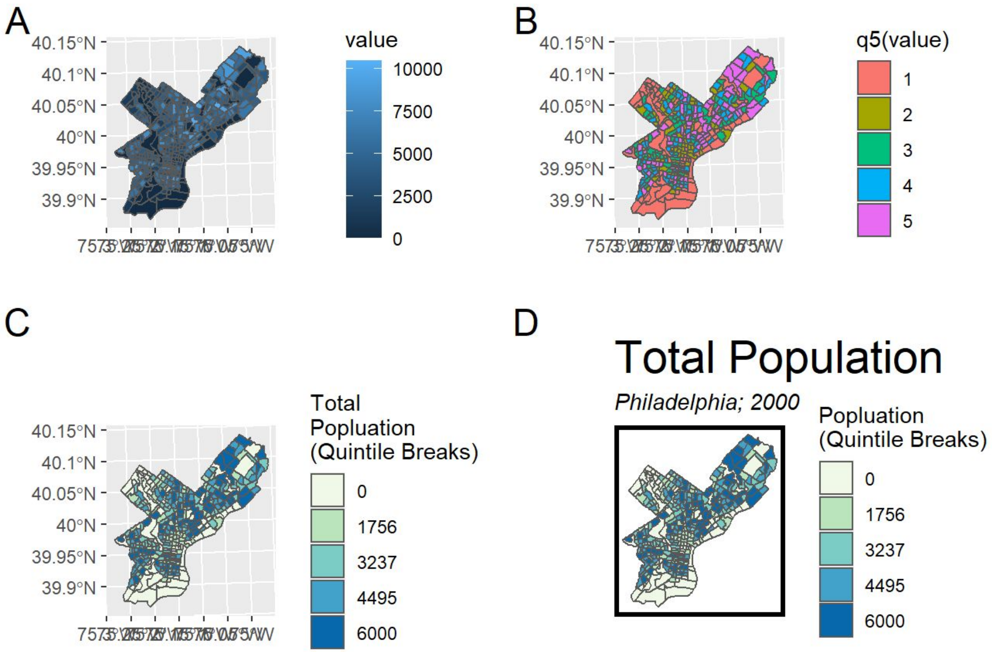


Figure 2.3

Download data from the web into R

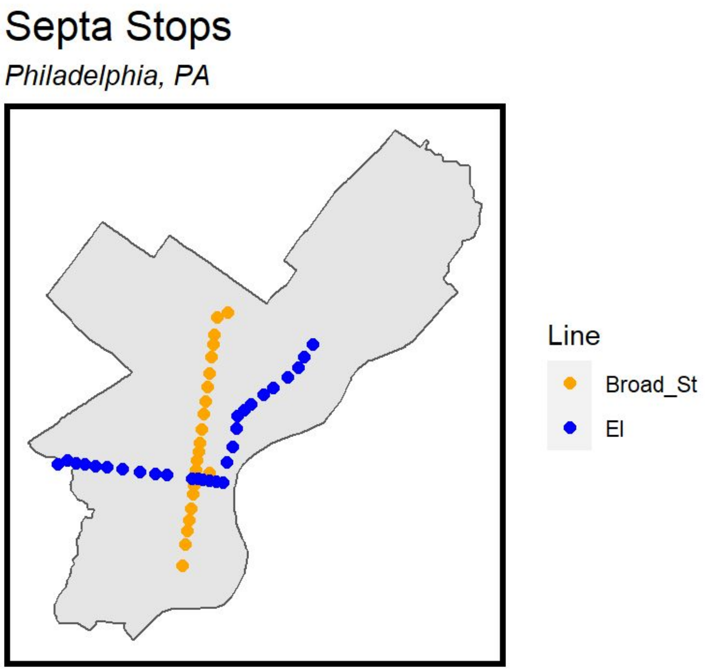
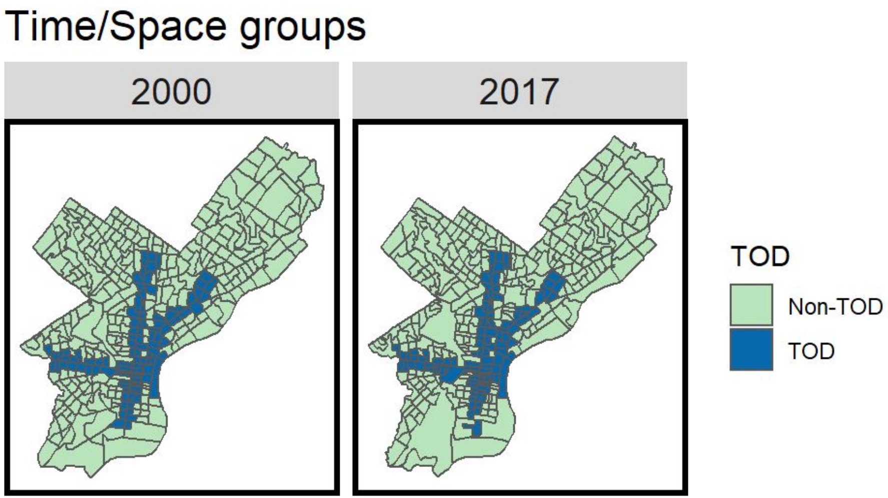
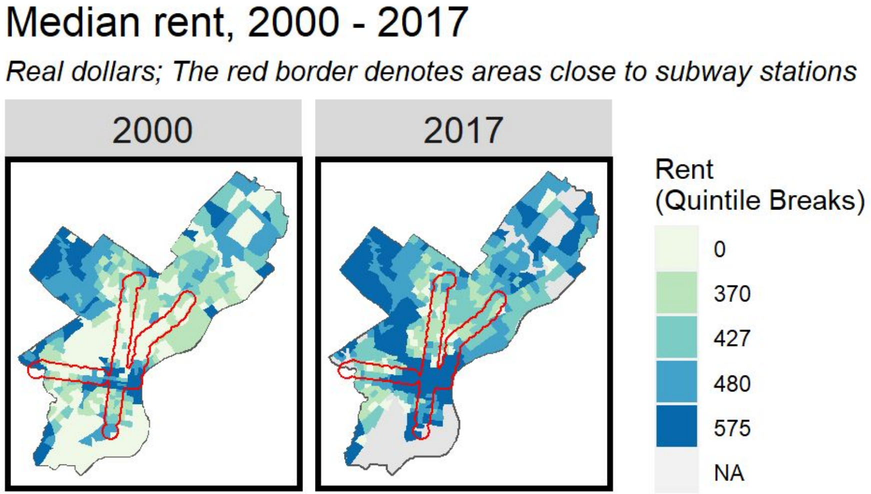


Figure 2.5

Relating tracts & Subway stops in space





Public Policy Analytics - Figure 2.8

Spatial scale: Relating tracts & Subway stops in space

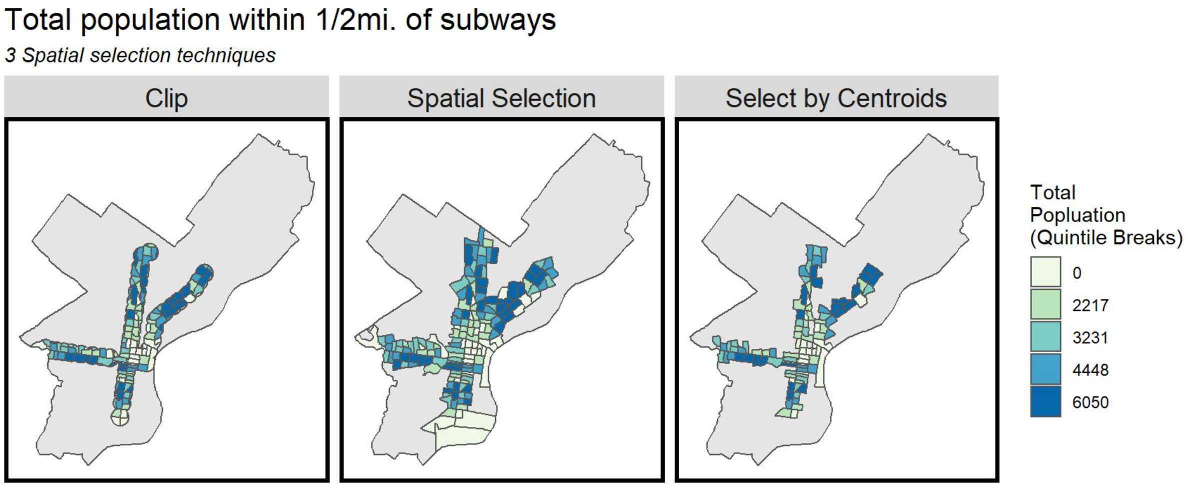


Figure 2.8

Relating tracts & Subway stops in space

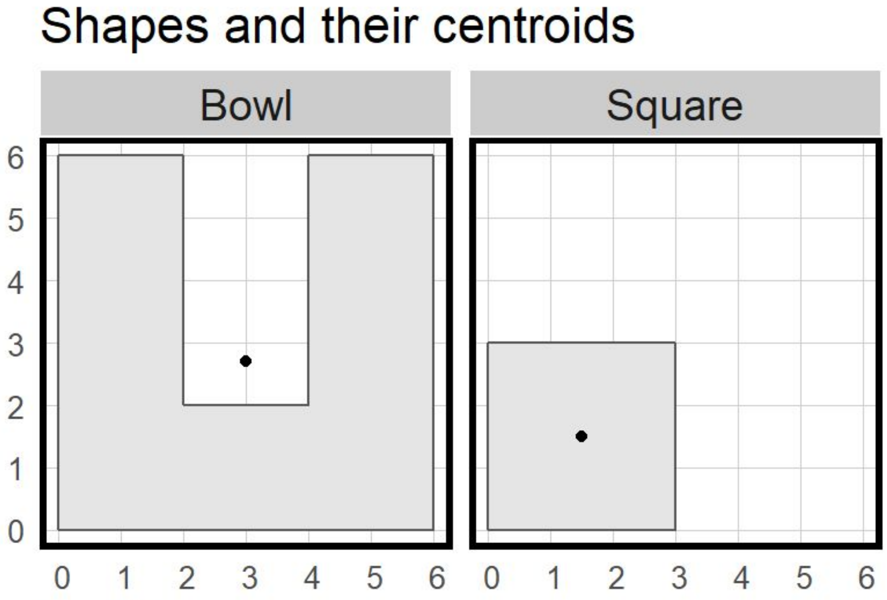


Figure 2.7

Now that we have Census data and have defined close, Let's look at some indicators - **Tables**

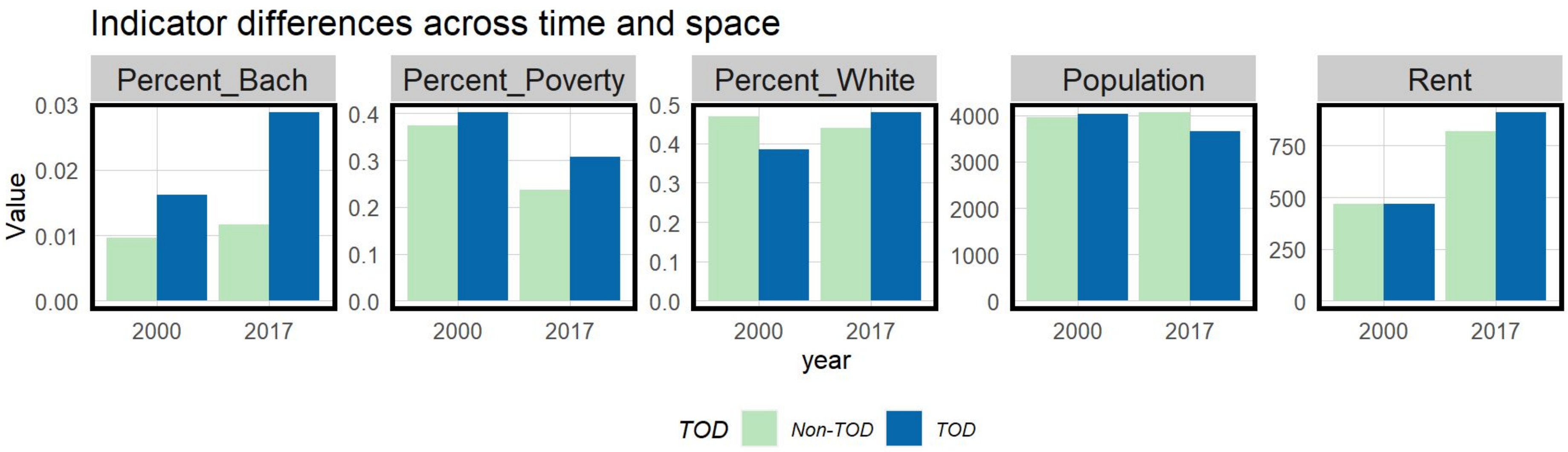
year	TOD	Rent	Population	Percent_White	Percent_Bach	Percent_Poverty
2000	Non-TOD	470.5458	3966.789	0.4695256	0.0096146	0.3735100
2000	TOD	469.8247	4030.742	0.3848745	0.0161826	0.4031254
2017	Non-TOD	821.1642	4073.547	0.4396967	0.0116228	0.2373258
2017	TOD	913.3750	3658.500	0.4803197	0.0288166	0.3080936

Now that we have Census data and have defined close, Let’s look at some indicators - **Tables**

Variable	2000: Non-TOD	2000: TOD	2017: Non-TOD	2017: TOD
Percent_Bach	0.01	0.02	0.01	0.03
Percent_Poverty	0.37	0.40	0.24	0.31
Percent_White	0.47	0.38	0.44	0.48
Population	3966.79	4030.74	4073.55	3658.50
Rent	470.55	469.82	821.16	913.38

Table 2.3

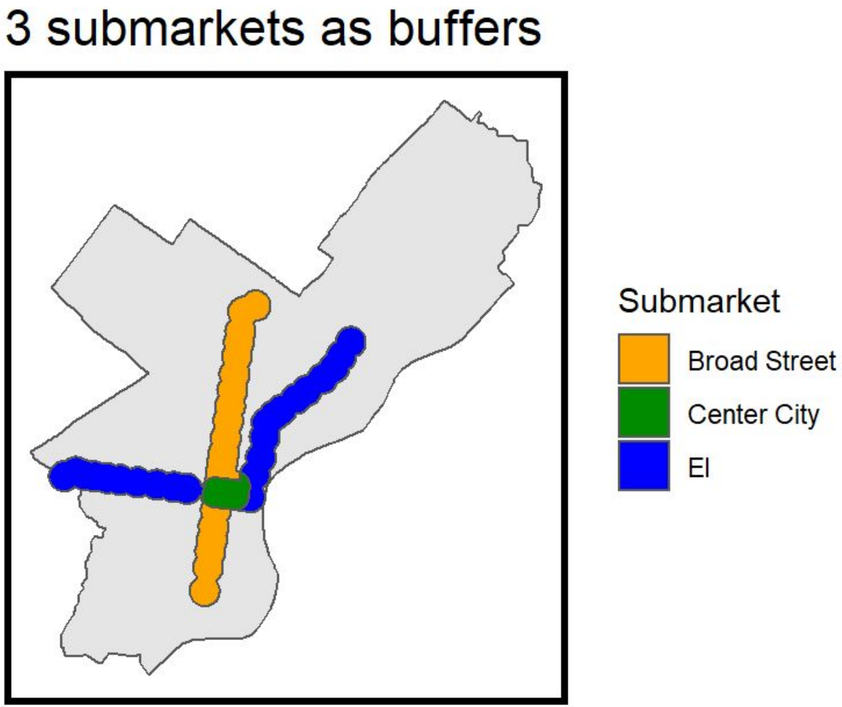
Now that we have Census data and have defined close, Let’s look at some indicators - **Plots**

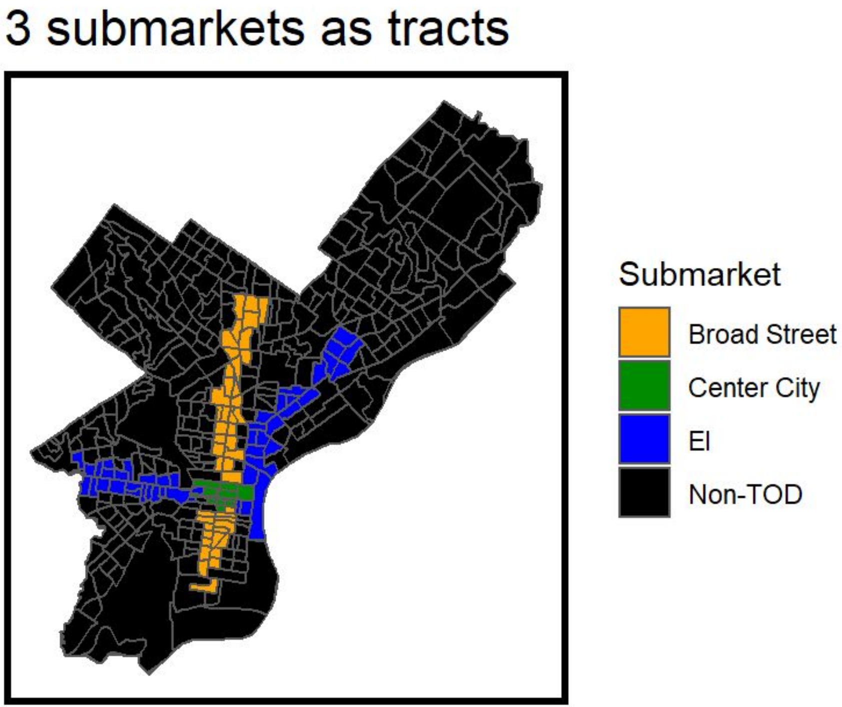


We have forgotten one key idea however…

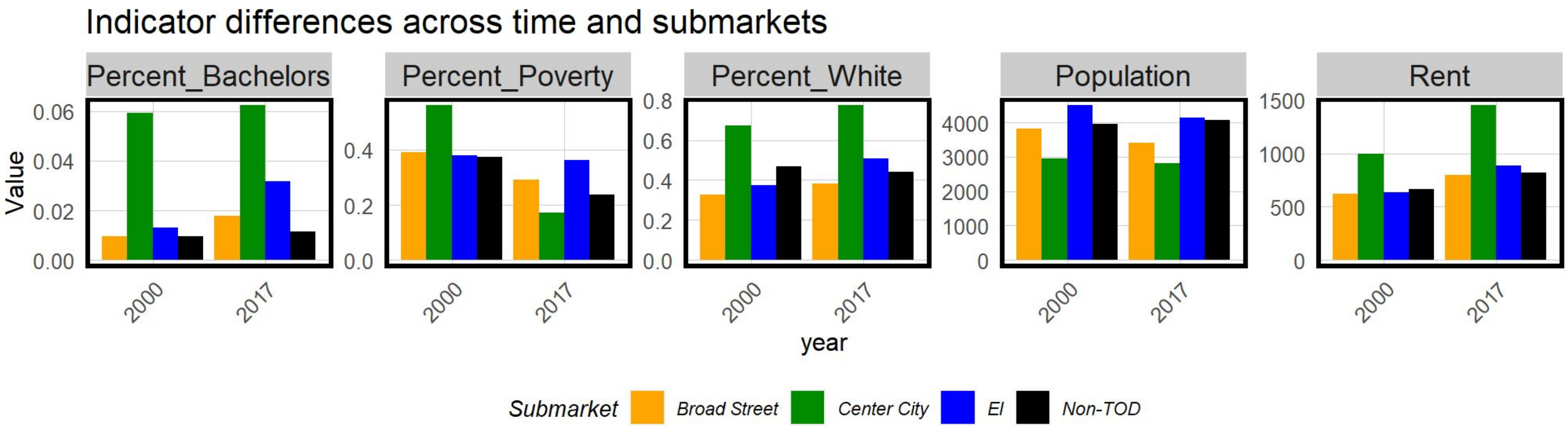
What is the **spatial process**?

Capturing 3 ‘submarkets’ of interest





Capturing 3 ‘submarkets’ of interest



Are Philadelphians willing to pay for TOD?

It could be that households are willing to pay more for transit amenities, or that they pay more for other amenities in neighborhoods that happen to also be transit-rich.

Although indicators enable the data scientist to simplify complex ideas, those ideas must be *interpreted* responsibly. This means acknowledging important assumptions in the data.

Assignment 2 - Study TOD in your city

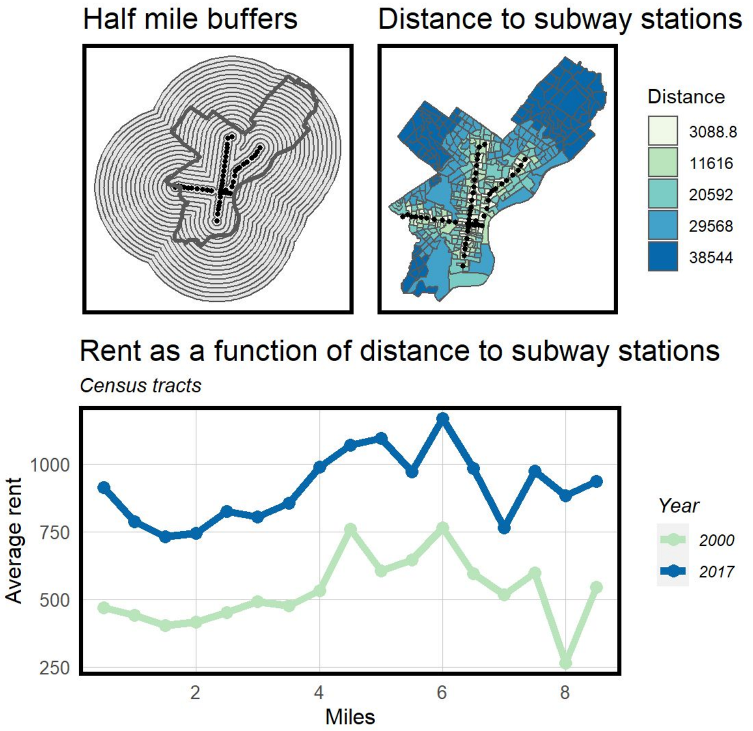


Figure xx WHAT IS IT?

Assignment 2 - Study TOD in your city

Recreate this analysis in a city of **your choosing** and prepare a policy brief (in R markdown) for local City Council representatives.

Do households value transit proximity?

Is this valuation even or uneven? How?

How certain can you be about your conclusions given some of the spatial biases we've discussed?

How can you collaborate with a teammate?

-Post on Canvas a City that has the requisite open data and solicit a partner.

-Project manage:

-Who will gather what data? What deliverables for data wrangling; visualization; markdown write-up.

-Use Github (more on this in lab)

(Don’t have an account? Sign up at github.com & DL Github Desktop)

Use data visualization to move the reader through the narrative.

Keep referring back to your research question throughout

DO NOT write a research paper!