Understanding the US Census and American Community Survey

Berkeley
UNIVERSITY OF CALIFORNIA

Introduction to Urban Data Analytics

Manuel Santana Palacios

May 29, 2020

Announcements

- Assignment 1
- Reading Response Questions
- Piazza
- Lectures video recordings
- Lab 1

Agenda for Today

Introduction to the Census & American Community Survey

- The 2020 Census
- History of the Census and ACS
- Geographic scales
- Race and ethnicity in Census

Lab 1: Excel Fundamentals Using ACS data

Reading Response Question

Eason Zhu

Singleton, 2018, "Urban Analytics" Chapter 6

The authors introduces the idea of identifying problematic data that might impede the research. In order to get a clean and accurate dataset, they argue that "identify missing data and erroneous attributes is a critical step in the analysis." It's important to distinguish what values are "missing" and what values are "zero", so how do researchers differentiate "missing" and "zero" data, and how do they appropriately manage these problematic data? Same questions applies to outliers as well, how do researchers assess what outliers to remove and what to keep in the data set? Do researchers consider sample size when they decide how many outliers or what outliers to drop?

United States Census 2020

Partners Educators News & Events Help

Search English

Get the Facts How To Respond Why Your Answers Matter Privacy and Security

RESPOND

Shape your future.

START HERE.

RESPOND

It's Quick and Easy

The 2020 Census takes just a few moments to complete.

It's Not Too Late

You can still complete the census today.

It's Safe and Secure

The U.S. Census Bureau keeps your answers safe and confidential.

Is this page helpful? ×
YES | NO

United States Census 2020

Partners Educators News & Events Help

Search English

Get the Facts How To Respond Why Your Answers Matter Privacy and Security

RESPOND

Shape your future.

How to Respond Online

How to Respond by Phone

How to Respond by Mail

Source: https://2020census.gov/en.html

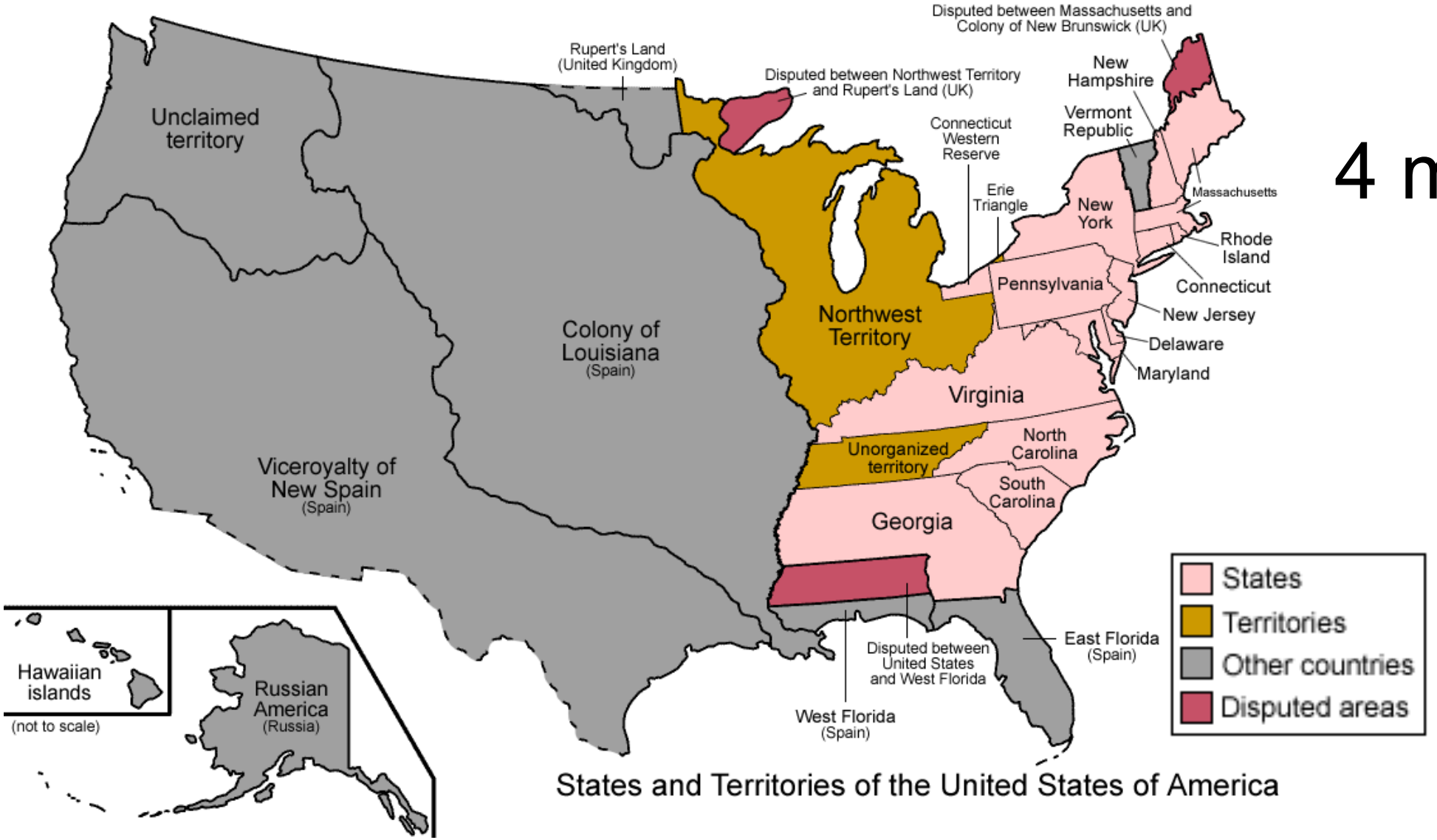
History

History of the Decennial Census

1930

Source: https://www.youtube.com/watch?v=IZMtEFjUgHM&feature=emb_rel_end

History of the Decennial Census



States and Territories of the United States of America

1790

4 million inh.

History of the Decennial Census

Source: https://www.youtube.com/watch?v=IZMtEFjUgHM&feature=emb_rel_end

The Census: A Cultural Document

1900 Included incidents of **deafness**, **blindness**, **insanity**, juvenile delinquency

1910 Added questions about **mines** and **quarries**. First time questions were asked about **nationality** or **mother tongue** of foreign-born persons and their parents

1940 First time information was collected on **housing**

1960 First time census forms were **mailed**, the birth of the "**long form**" (SF3 and SF4)

1980 First time questions were asked about **Hispanic origin** of 100% of the population

1990 First time a household could indicate "**same-sex**" partnership

2000 First time an individual could indicate **mixed-race** identity

The ‘Short’ and ‘Long’ Form
1970, 1990, 1990 & 2000

The “Short Form”
[Summary File 1]

- 100 % sample

The “long form”
[Summary File 3]

- about 1 in 6 households

The ‘Short’ and ‘Long’ Form
2000 Census

The “Short Form”
[Summary File 1]

- **100 % sample**
- Count of people, households, housing units
- Age, sex, race/ethnicity,
- Owner/renter status
- Family composition

The “long form”
[Summary File 3]

- **About 1 in 6 households**
- Sample of about 19 million housing units
- Social, economic and housing characteristics

2010: The Beginning of the ACS

Decennial Census
The “Short Form”

~~The “long form”~~
American Community Survey

2010: The Beginning of the ACS

Decennial Census
The "Short Form"

~~The "long form"~~

American Community Survey

- 'Rolling' sample
- 3 million housing units each year
- Sample size much smaller than 'long form'
- **MoE** become important

2010: The Beginning of the ACS

Decennial Census
The "Short Form"

~~The "long form"~~
American Community? Survey

Defining Community

- **What** or **who** is the community? People, houses, activity?
- When we're working with data, what is the **boundary**? Political or census boundaries may not align with "places" and their social meanings
- **Zip codes** increasingly more common with administrative or private sources of data (because of link to street address)
- Most researchers use **one or more** census **tracts** to define neighborhoods

Balance between recognizing the complexity of a community and the need to “make plans” or “take action”

The American Community Survey

1-year estimates	3-year estimates*	5-year estimates
12 months of collected data *Example:* 2018 ACS 1-year estimates *Date collected between:* January 1, 2018 and December 31, 2018	36 months of collected data *Example:* 2011-2013 ACS 3-year estimates *Date collected between:* January 1, 2011 and December 31, 2013	60 months of collected data *Example:* 2014-2018 ACS 5-year estimates *Date collected between:* January 1, 2014 and December 31, 2018
Data for areas with populations of 65,000+	Data for areas with populations of 20,000+	Data for all areas

https://www.census.gov/programs-surveys/acs/guidance/estimates.html

The American Community Survey

1-year estimates	3-year estimates*	5-year estimates
12 months of collected data *Example:* 2018 ACS 1-year estimates *Date collected between:* January 1, 2018 and December 31, 2018	36 months of collected data *Example:* 2011-2013 ACS 3-year estimates *Date collected between:* January 1, 2011 and December 31, 2013	60 months of collected data *Example:* 2014-2018 ACS 5-year estimates *Date collected between:* January 1, 2014 and December 31, 2018
states, big counties 65,000+	Data for areas with populations of 20,000+	census tracts, block groups

https://www.census.gov/programs-surveys/acs/guidance/estimates.html

The American Community Survey

	1-year estimates	3-year estimates*	5-year estimates
	12 months of collected data *Example:* 2018 ACS 1-year estimates *Date collected between:* January 1, 2018 and December 31, 2018	36 months of collected data *Example:* 2011-2013 ACS 3-year estimates *Date collected between:* January 1, 2011 and December 31, 2013	60 months of collected data *Example:* 2014-2018 ACS 5-year estimates *Date collected between:* January 1, 2014 and December 31, 2018
Units of analysis			
Aggregate (places)	states, big counties 65,000+	Data for areas with populations of 20,000+	census tracts, block groups

https://www.census.gov/programs-surveys/acs/guidance/estimates.html

Units of analysis

Commuting in Berkeley Analyses

Aggregate

Tract	Attending college	Commutes by bike
Tract 4222	0.25	0.14
Tract 4223	0.13	0.08
Tract 4224	0.55	0.52

Disaggregate

Person	Attending college	Commutes by bike
125154324	1	0
125154325	0	1
125154326	1	1

Units of analysis

Commuting in Berkeley

Aggregate

Tract	Attending college	Commutes by bike
Tract 4222	0.25	0.14
Tract 4223	0.13	0.08
Tract 4224	0.55	0.52

Disaggregate

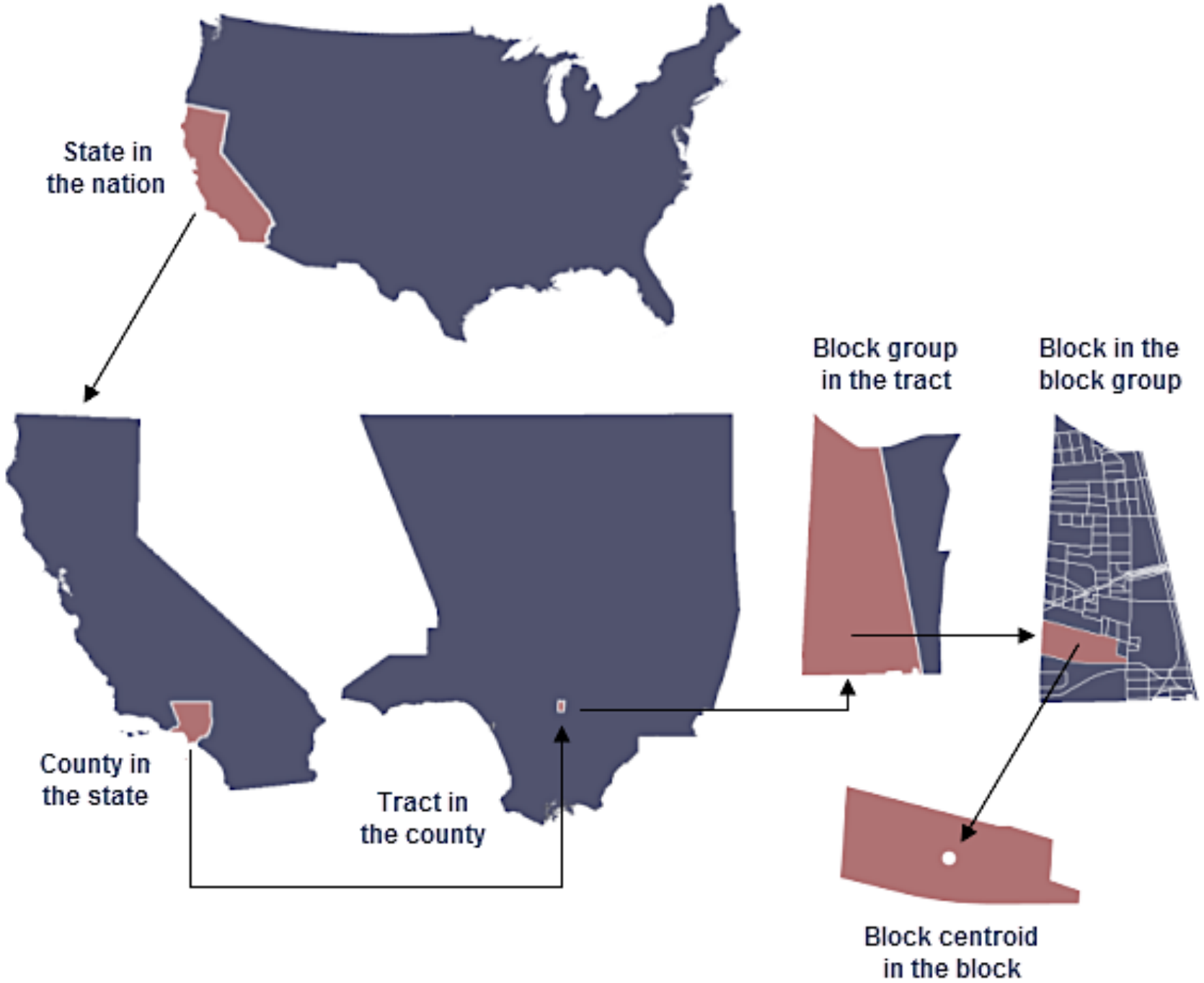
Person	Attending college	Commutes by bike
125154324	1	0
125154325	0	1
125154326	1	1

Beware of the **ecological fallacy**!

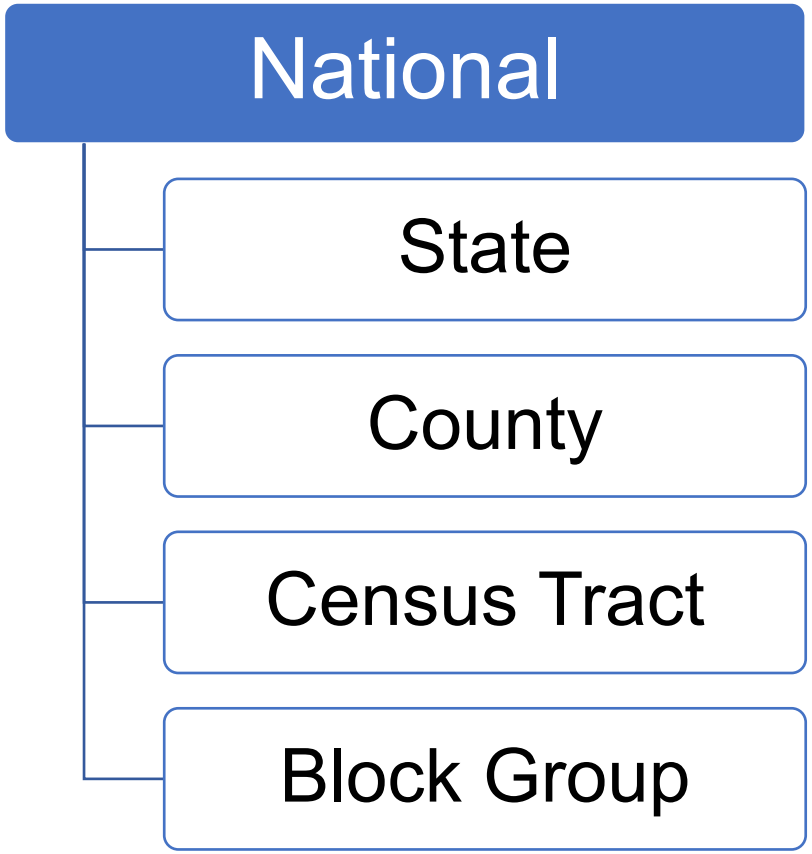
making inferences about individuals from group data

Geographic Scales

Key Census Building Blocks

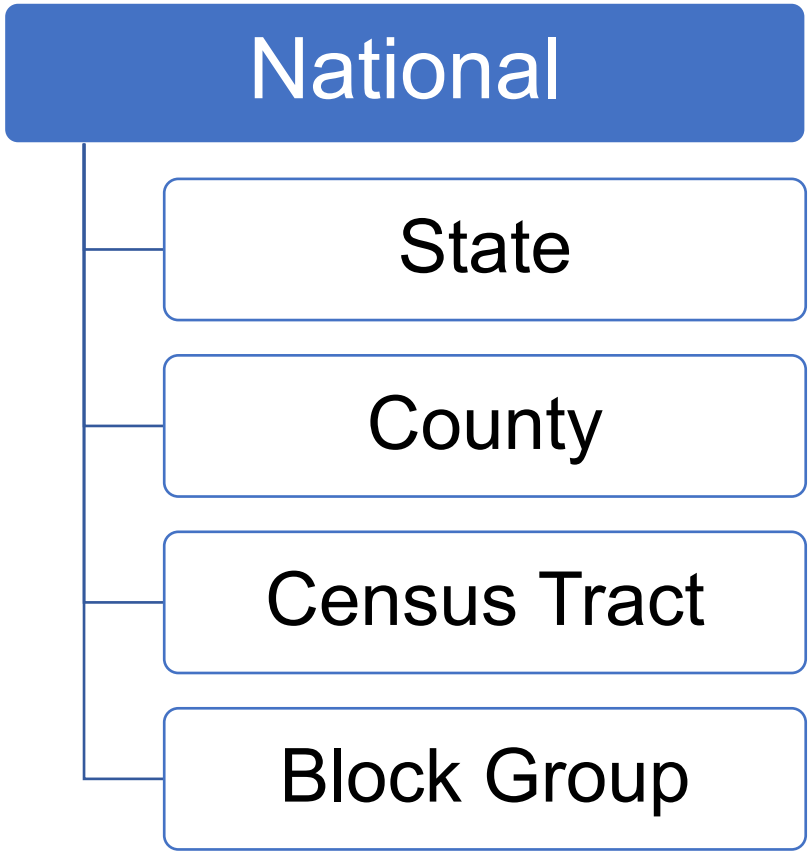


Key Census Building Blocks



Nested geographies identified by a Federal Information Processing Standards (**FIPS**) code

Key Census Building Blocks



06 - California

001 - Alameda County

401100 - in Temescal**

1 - Block Group

Key Census Building Blocks

FIPS
060014011001

06 - California

001 - Alameda County

401100 - in Temescal**

1 - Block Group

Key Census Building Blocks

FIPS

06001401100 1

California | Alameda County | Tract | Block Group

Temescal**

Race and Ethnicity

Race/Ethnicity

Brazil 2000 Census

Choose your race:

- ☐ White - branca
- ☐ Black - preta
- ☐ Yellow - amarela
- ☐ Brown - parda
- ☐ Native, aboriginal - indigena
- ☐ Undeclared

England 2001 Census

What is your ethnic group?

A) White:

☐ British ☐ Irish ☐ Any other white background

B) Mixed

☐ White and Black Caribbean ☐ White and Black African

☐ White and Asian ☐ Any other Mixed background

C) Asian or Asian British

☐ Indian ☐ Pakistani

☐ Bangleshi ☐ Any other Asian background

D) Black or Black British

☐ Caribbean ☐ African ☐ Any other black background

E) Chinese or other ethnic group

☐ Chinese ☐ Any other

Race/Ethnicity is a **Social Construct**

Brazil 2000 Census

Choose your race:

- ☐ White - branca
- ☐ Black - preta
- ☐ Yellow - amarela
- ☐ Brown - parda
- ☐ Native, aboriginal - indigena
- ☐ Undeclared

England 2001 Census

What is your ethnic group?

A) White:

☐ British ☐ Irish ☐ Any other white background

B) Mixed

☐ White and Black Caribbean ☐ White and Black African

☐ White and Asian ☐ Any other Mixed background

C) Asian or Asian British

☐ Indian ☐ Pakistani

☐ Bangleshi ☐ Any other Asian background

D) Black or Black British

☐ Caribbean ☐ African ☐ Any other black background

E) Chinese or other ethnic group

☐ Chinese ☐ Any other

➜ **NOTE: Please answer BOTH Question 8 about Hispanic origin and Question 9 about race. For this census, Hispanic origins are not races.**

8. Is Person 1 of Hispanic, Latino, or Spanish origin?

- ☐ **No**, not of Hispanic, Latino, or Spanish origin
- ☐ Yes, Mexican, Mexican Am., Chicano
- ☐ Yes, Puerto Rican
- ☐ Yes, Cuban
- ☐ Yes, another Hispanic, Latino, or Spanish origin – *Print, for example, Salvadoran, Dominican, Colombian, Guatemalan, Spaniard, Ecuadorian, etc.*

9. What is Person 1's race?

*Mark [X] one or more boxes **AND** print origins.*

- [ ] White – *Print, for example, German, Irish, English, Italian, Lebanese, Egyptian, etc.*

- [ ] Black or African Am. – *Print, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc.*

- [ ] American Indian or Alaska Native – *Print name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Traditional Government, Nome Eskimo Community, etc.*

- [ ] Chinese
- [ ] Filipino
- [ ] Asian Indian
- [ ] Vietnamese
- [ ] Korean
- [ ] Japanese
- [ ] Native Hawaiian
- [ ] Samoan
- [ ] Chamorro
- [ ] Other Asian – *Print, for example, Pakistani, Cambodian, Hmong, etc.*
- [ ] Other Pacific Islander – *Print, for example, Tongan, Fijian, Marshallese, etc.*

- [ ] Some other race – *Print race or origin.*

The Myth of the White Minority

Reading Response Question

Annie Ouyang

Alba, 2015, The Myth of a White Minority

The article has raised the question of what should be the definition of "minority" in America. With the increase in immigration as well as cultural exchanges, is it still reasonable to group people based on their skin color and appearance? How do we account for the many multi-generational families who no longer connect with the heritage of their ancestors in terms of language and art? The author argues that we cannot be completely a post-racial society, as we "cannot abandon ethno-racial categories. They register legacies of slavery, conquest and oppression that have enduring effects." Nevertheless, how important is race in telling one's story and beyond race, are there other means to capture the essence of one's identity?

Race/Ethnicity is a **Social Construct**

What's wrong here?

Table 1 - Population Demographics, Alameda County

	Population	Percent
White	517,76	33.7
Black or African American	181,315	11.8
American Indian and Alaska Native	4,484	0.3
Asian	408,556	28.6
Native Hawaiian and Other Pacific Islander	12,230	0.8

Source: ACS 2013

Race/Ethnicity is a **Social Construct**

What's wrong here?

Table 1 - Population Demographics, Alameda County

	Population	Percent
White	517,76	33.7
Black or African American	181,315	11.8
American Indian and Alaska Native	4,484	0.3
Asian	408,556	28.6
Native Hawaiian and Other Pacific Islander	12,230	0.8

Source: ACS 2013

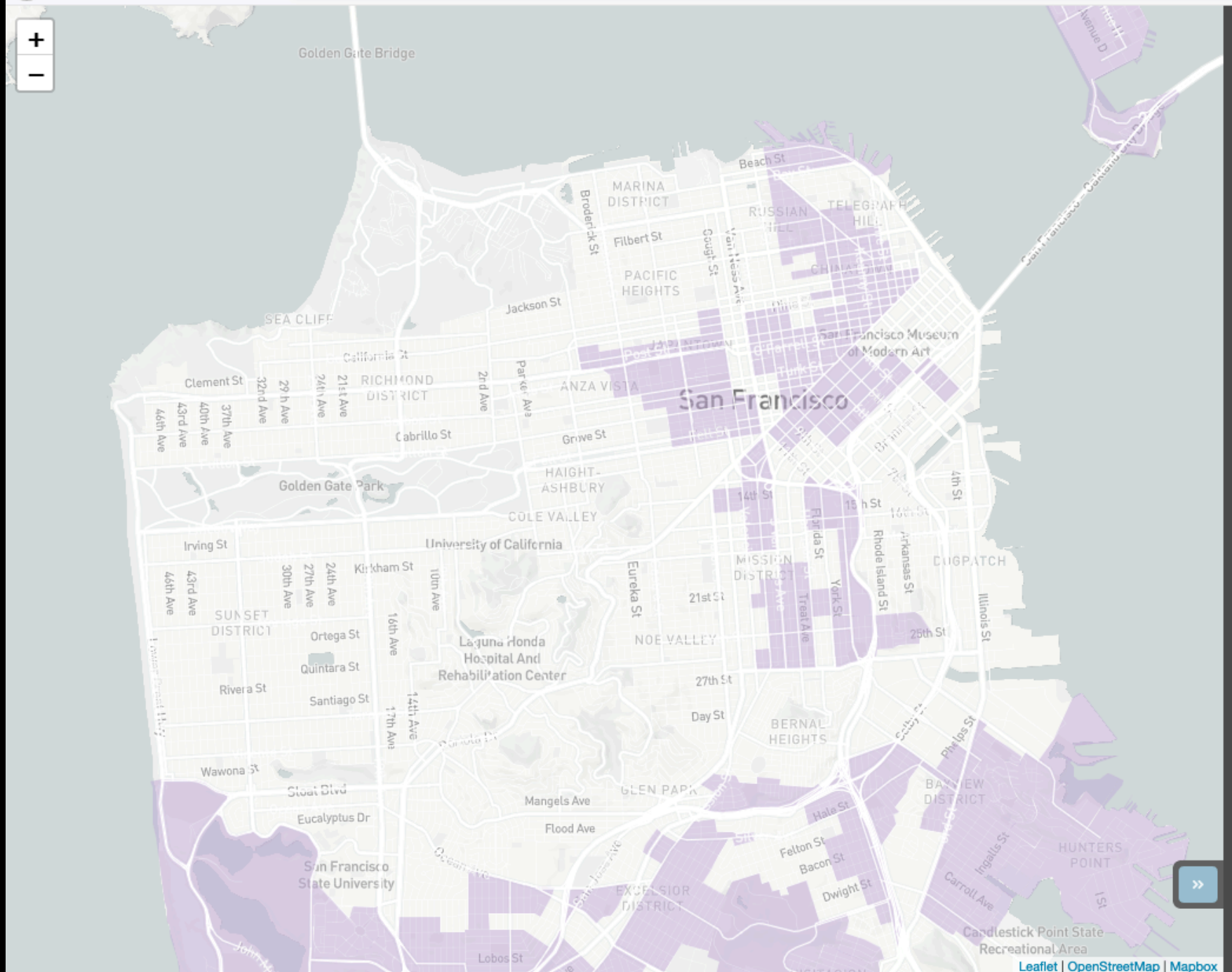
Race/Ethnicity is a **Social Construct**

Table 1 - Population Demographics, Alameda County

	Population	Percent
Non-Hispanic White	517,764 ~~700,110~~	33.7 ~~45.6~~
Black or African American	181,315 ~~185,467~~	11.8 ~~12.1~~
American Indian and Alaska Native	4,484 ~~8,919~~	0.3 ~~0.6~~
Hispanic	345,847	22.5
Asian	408,556 ~~411,240~~	28.6 ~~26.8~~
Native Hawaiian and Other Pacific Islander	12,230 ~~12,579~~	0.8

Source: ACS 2013, Table B03002 '**Hispanic or Latino Origin by Race**'

Communities of Concern



Communities of Concern

What is this? | Learn More

CHOOSE MEASURE

None | Minority | Low Income | Elderly

Disability | Low English Prof. | Zero-Veh HH

Single Parent | Rent Burdened

MAP LAYERS

Add additional geographic data to the map.

- High Injury Network
- Supervisorial District Boundaries
- Major Parks

YOUR FEEDBACK

Please provide feedback. What do you think about this map? (800 maximum characters)

This site is protected by reCAPTCHA and the Google Privacy Policy and Terms of Service apply.

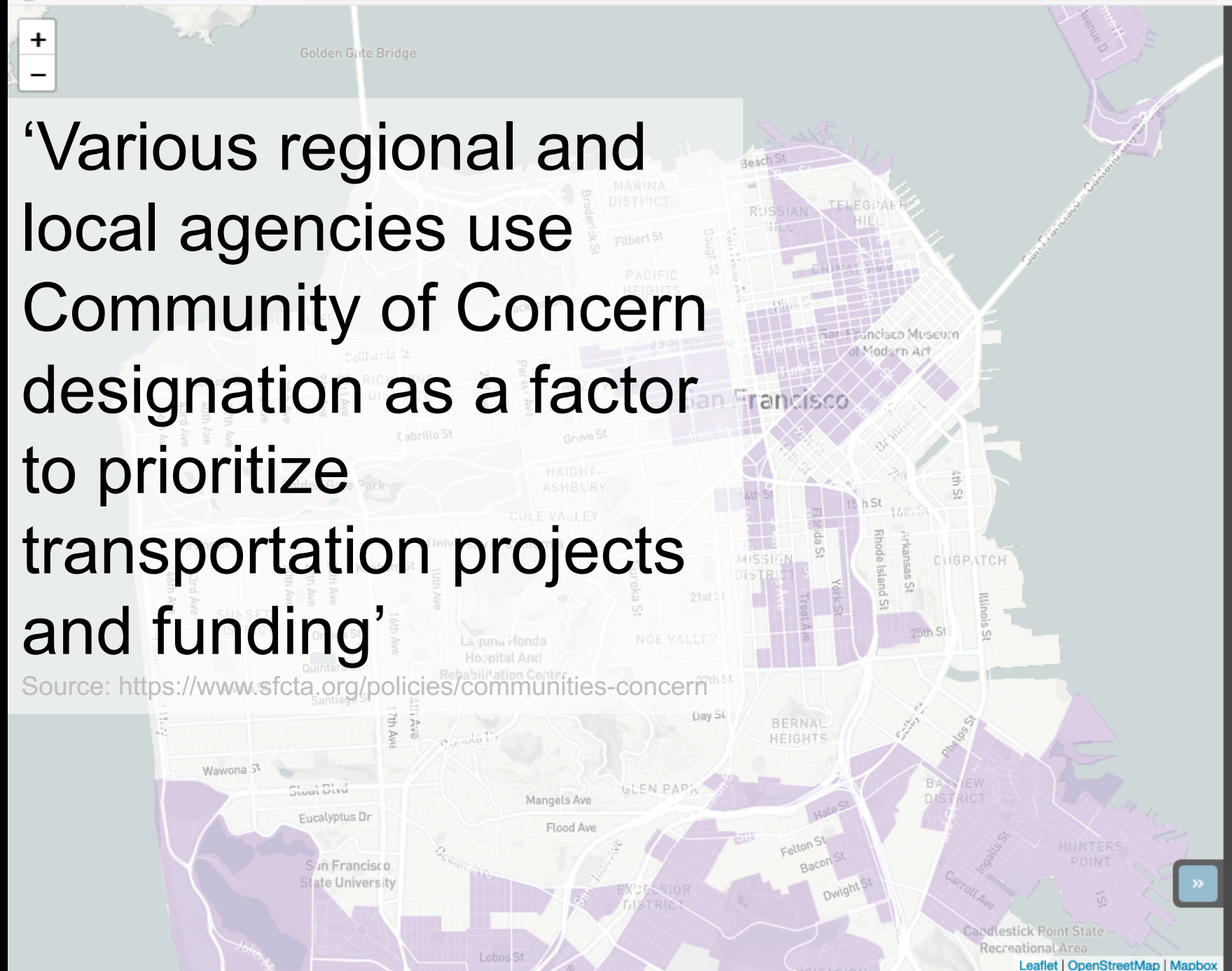
Submit

San Francisco County Transportation Authority

Communities of Concern

‘Various regional and local agencies use Community of Concern designation as a factor to prioritize transportation projects and funding’

Source: https://www.sfcta.org/policies/communities-concern



Lab 1 – Excel Fun with ACS Data