



Using Census Bureau data products with R

More info: micom-hub.github.io

Outline

- Slides: ~15 minutes
- Coding: the rest of the time
 - Querying the Census Bureau's API using `tidycensus`
 - Querying the Census Bureau's TIGERweb API using `tigris`
 - Putting these assets together with our fake MDSS data using `sf`
- If there's time...
 - Applying basemaps with `ggmap` and OpenStreetMap
 - Looking at interactive maps with `leaflet`

The US Census Bureau

- Two main data products
 - As far as today is concerned, anyway
- The Decennial Census
 - Actually, we won't talk about this too much, but everyone's heard of it
- The American Community Survey
 - Often what you *actually* want when you're thinking of using Census data
- They have *tons* of other surveys, too
- Beyond that, they're actually just one of **13** principal statistical agencies in the federal government



What's the difference between the Census and the ACS?

- The decennial census is mandated by the Constitution to be conducted as a literal count of people, everywhere in the country
 - This isn't really how statisticians come up with estimates of population sizes anymore
- The American Community Survey is the Census Bureau's response to the statistical shortcomings of the literal-minded Census
 - Estimates (try to) account for underreporting
 - They come with margins of error (90%, 95%, 99%)



versus



So, what's in the American Community Survey?

Tons of stuff. [Let's just look at the variables table!](#)

Census Data API: Variables in /data/2024/acs/1/variables							
Name	Label	Concept	Required	Attributes	Limit	Predicate Type	Group
AIANHH	American Indian Area/Alaska Native Area/Hawaiian Home Land		not required		0	(not a predicate)	N/A
ANRC	Alaska Native Regional Corporation		not required		0	(not a predicate)	N/A
B01001_001E	Estimate!!Total:	Sex by Age	not required	B01001_001EA, B01001_001M, B01001_001MA	0	int	B01001
B01001_002E	Estimate!!Total:!!Male:	Sex by Age	not required	B01001_002EA, B01001_002M, B01001_002MA	0	int	B01001
B01001_003E	Estimate!!Total:!!Male:!!Under 5 years	Sex by Age	not required	B01001_003EA, B01001_003M, B01001_003MA	0	int	B01001
B01001_004E	Estimate!!Total:!!Male:!!5 to 9 years	Sex by Age	not required	B01001_004EA, B01001_004M, B01001_004MA	0	int	B01001
B01001_005E	Estimate!!Total:!!Male:!!10 to 14 years	Sex by Age	not required	B01001_005EA, B01001_005M, B01001_005MA	0	int	B01001
B01001_006E	Estimate!!Total:!!Male:!!15 to 17 years	Sex by Age	not required	B01001_006EA, B01001_006M, B01001_006MA	0	int	B01001
B01001_007E	Estimate!!Total:!!Male:!!18 and 19 years	Sex by Age	not required	B01001_007EA, B01001_007M, B01001_007MA	0	int	B01001
B01001_008E	Estimate!!Total:!!Male:!!20 years	Sex by Age	not required	B01001_008EA, B01001_008M, B01001_008MA	0	int	B01001
B01001_009E	Estimate!!Total:!!Male:!!21 years	Sex by Age	not required	B01001_009EA, B01001_009M, B01001_009MA	0	int	B01001
B01001_010E	Estimate!!Total:!!Male:!!22 to 24 years	Sex by Age	not required	B01001_010EA, B01001_010M, B01001_010MA	0	int	B01001
B01001_011E	Estimate!!Total:!!Male:!!25 to 29 years	Sex by Age	not required	B01001_011EA, B01001_011M, B01001_011MA	0	int	B01001
B01001_012E	Estimate!!Total:!!Male:!!30 to 34 years	Sex by Age	not required	B01001_012EA, B01001_012M,	0	int	B01001

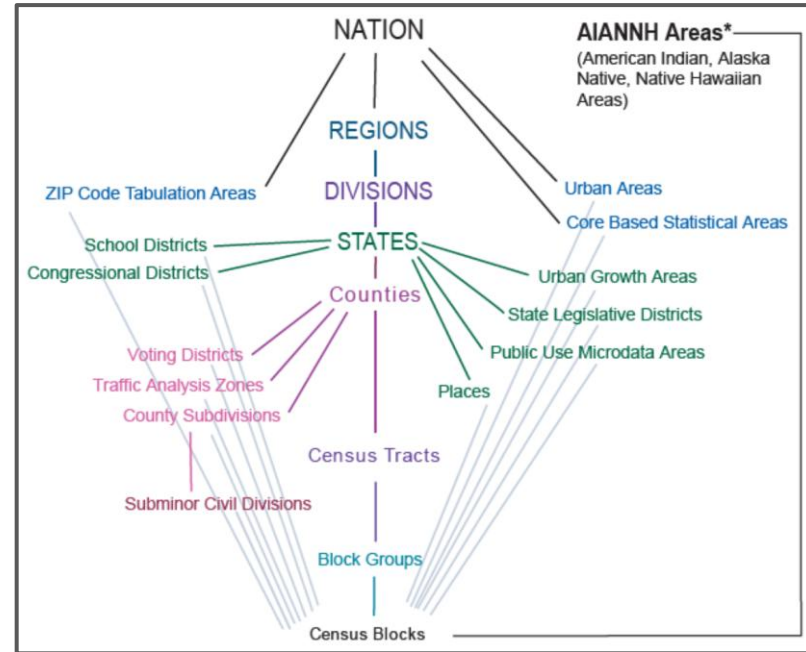
The ACS is split into two data products: 1-Year and 5-Year estimates

The census is really expensive! That's the other reason the feds do the American Community Survey.

- The ACS is distributed to ~3% of the population annually, about 3.5 million households
- 5-Year estimates are made for all Census geographies, down to the block (usually < 1,000 people)
 - Most recent available set: 2020-2024
- Annual (1-Year) estimates are made for all geographies with at least 65,000 people
 - Most recent available set: 2024
 - Medium-sized counties and above, in terms of on-spine geographies.
 - Wait... what are "on-spine geographies"??

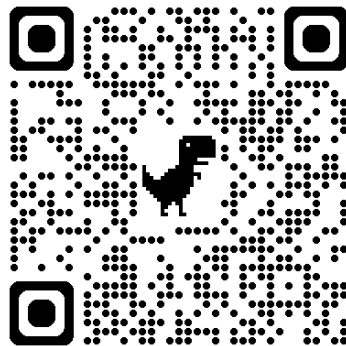
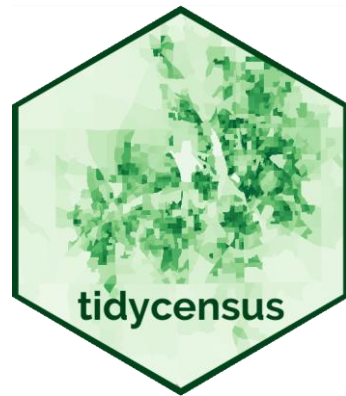
Statistical geographies vs legal/administrative geographies

- Legal/administrative geographies
 - States, counties, ZIP codes, cities, school districts, voting precincts, **health department jurisdictions**
- Statistical geographies
 - Census tracts, block groups, metropolitan statistical areas (MSAs), public-use microdata areas (PUMAs)
- ACS estimates are designed to be robust along “on-spine” geographies
 - Hierarchical & geographically nested
 - Nation → state → county → tract → block



The Census API

- Provides programmatic access to all Census data products - for free!
 - No promises, but the API kept working for me during the government shutdown about a year ago.
- Easiest access requires each user make an API key
 - Again, free!
 - On a safe **.gov** domain
 - You can do it on your phone
 - Takes two seconds
 - Dead easy
 - Please do it :-)
 - Then check your email after a minute
- We R users are lucky to have **tidycensus**!



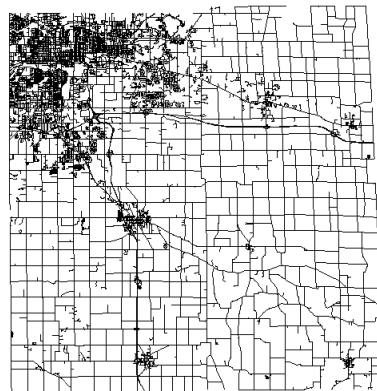
[Census API
key signup](#)

The Census API: geography

- The Census Bureau also runs TIGERweb, a geospatial webservice with an API
 - Counties
 - Cities
 - ZIP codes
 - Tracts
 - Block groups
 - School districts
 - Tribal geographies
 - Rivers & lakes
 - Roads
 - It's got the stinkin' *railroads!*
- We R users are *also* lucky to have **tigris**!



Ingham Co. roads, from 'tigris'



Let's get into it!