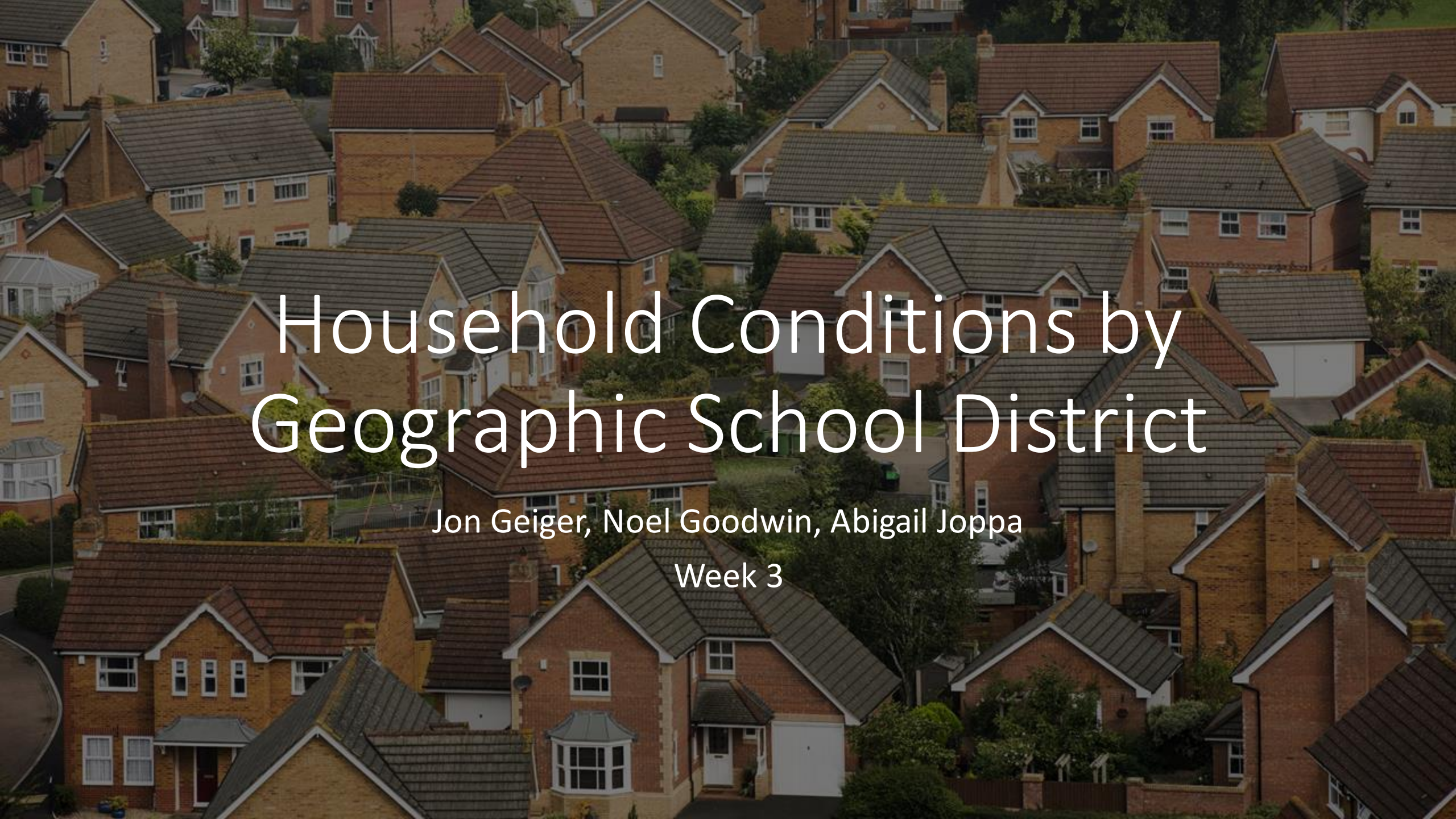


Household Conditions by Geographic School District

Jon Geiger, Noel Goodwin, Abigail Joppa

Week 3

Review

Correlations between Measures of Vulnerability across PUMAs

Lack of internet or computer access is most correlated with poverty at the PUMA level

	Is in poverty	Is linguistically isolated	Has disability	Is in vulnerable economic sector	Has single parent	Is in crowded conditions	Lacks computer or broadband
Is in poverty	1.00						
Is linguistically isolated	0.38	1.00					
Has disability	0.46	-0.20	1.00				
Is in vulnerable economic sector	0.52	0.33	0.23	1.00			
Has single parent	0.74	0.13	0.45	0.32	1.00		
Is in crowded conditions	0.53	0.71	-0.09	0.44	0.16	1.00	
Lacks computer or broadband	0.80	0.26	0.39	0.47	0.50	0.45	1.00

URBAN INSTITUTE

Source: Urban Institute analysis of 2014–18 American Community Survey data.
Note: PUMA = public-used microdata area.

One row is one school district

Last Week's Challenges

- Margins of Error (Jon)
- Variable Definitions (Noel)
- External Data Sets (Noel + Abigail)
- Organization and Management (All)

Margins of Error

A Peek at the Data

```
## # A tibble: 5 × 4
##   school_ID children pct_SP SP_MOE
##   <chr>         <dbl>  <dbl> <chr>
## 1 00001           985 0.0490 0%-10%
## 2 00003           292 0.102  3%-17%
## 3 00005          4591 0.353  26%-44%
## 4 00006          8299 0.295  24%-35%
## 5 00007         15397 0.208  17%-25%
```

Relevant Equations

$$\mathrm{SE}_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\mathrm{MOE}_{\hat{p}} = \hat{p} \pm z^{*}(\mathrm{SE}_{\hat{p}})$$

Why doesn't this work?

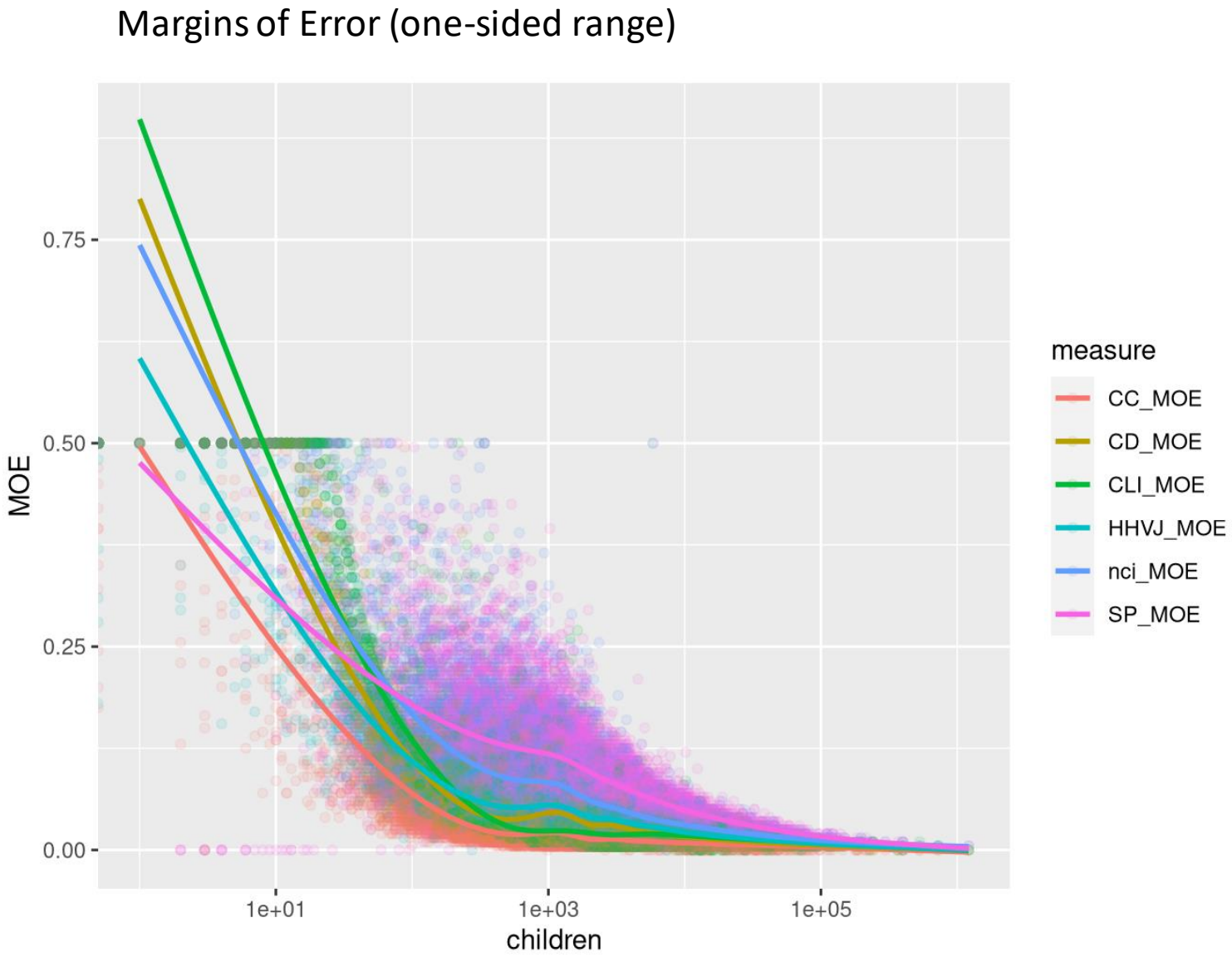
- "The ACS uses a successive differences replication (SDR) variance estimation methodology to derive the MOEs ... Taking the square root of the variance produces the standard error. The 90% confidence level MOE is equal to the standard error multiplied by 1.645."*

- Their standard errors are more complicated than the simple equation.

- What's *n* ?

* https://www2.census.gov/programs-surveys/acs/replicate_estimates/2018/documentation/5-year/2014-2018_Variance_Replicate_Tables_Documentation.pdf

Okay, so how *do* the margins of error behave?



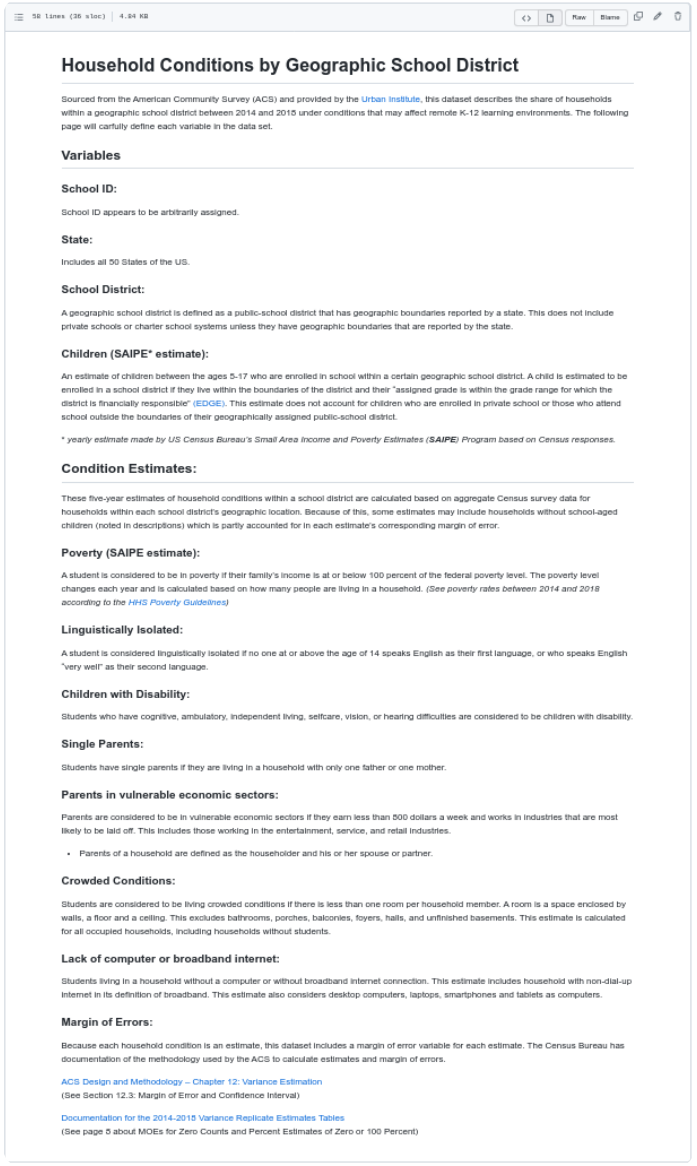
Variable Definitions

Why Estimates?

- Five-year estimates
- Census survey data
- Households within each school district’s geographic location

More Information?

https://github.com/noelgoodwin/household-conditions#readme



Household Conditions by Geographic School District

Sourced from the American Community Survey (ACS) and provided by the Urban Institute, this dataset describes the share of households within a geographic school district between 2014 and 2015 under conditions that may affect remote K-12 learning environments. The following page will carfully define each variable in the data set.

Variables

School ID:

School ID appears to be arbitrarily assigned.

State:

Includes all 50 States of the US.

School District:

A geographic school district is defined as a public-school district that has geographic boundaries reported by a state. This does not include private schools or charter school systems unless they have geographic boundaries that are reported by the state.

Children (SAIPE* estimate):

An estimate of children between the ages 5-17 who are enrolled in school within a certain geographic school district. A child is estimated to be enrolled in a school district if they live within the boundaries of the district and their "assigned grade is within the grade range for which the district is financially responsible" (EDGE). This estimate does not account for children who are enrolled in private school or those who attend school outside the boundaries of their geographically assigned public-school district.

** yearly estimate made by US Census Bureau's Small Area Income and Poverty Estimates (**SAIPE**) Program based on Census responses.*

Condition Estimates:

These five-year estimates of household conditions within a school district are calculated based on aggregate Census survey data for households within each school district's geographic location. Because of this, some estimates may include households without school-aged children (noted in descriptions) which is partly accounted for in each estimate's corresponding margin of error.

Poverty (SAIPE estimate):

A student is considered to be in poverty if their family's income is at or below 100 percent of the federal poverty level. The poverty level changes each year and is calculated based on how many people are living in a household. *(See poverty rates between 2014 and 2018 according to the HHS Poverty Guidelines)*

Linguistically Isolated:

A student is considered linguistically isolated if no one at or above the age of 14 speaks English as their first language, or who speaks English "very well" as their second language.

Children with Disability:

Students who have cognitive, ambulatory, independent living, selfcare, vision, or hearing difficulties are considered to be children with disability.

Single Parents:

Students have single parents if they are living in a household with only one father or one mother.

Parents in vulnerable economic sectors:

Parents are considered to be in vulnerable economic sectors if they earn less than 500 dollars a week and works in industries that are most likely to be laid off. This includes those working in the entertainment, service, and retail industries.

- Parents of a household are defined as the householder and his or her spouse or partner.

Crowded Conditions:

Students are considered to be living crowded conditions if there is less than one room per household member. A room is a space enclosed by walls, a floor and a ceiling. This excludes bathrooms, porches, balconies, foyers, halls, and unfinished basements. This estimate is calculated for all occupied households, including households without students.

Lack of computer or broadband internet:

Students living in a household without a computer or without broadband internet connection. This estimate includes household with non-dial-up internet in its definition of broadband. This estimate also considers desktop computers, laptops, smartphones and tablets as computers.

Margin of Errors:

Because each household condition is an estimate, this dataset includes a margin of error variable for each estimate. The Census Bureau has documentation of the methodology used by the ACS to calculate estimates and margin of errors.

ACS Design and Methodology – Chapter 12: Variance Estimation
(See Section 12.3: Margin of Error and Confidence Interval)

Documentation for the 2014-2018 Variance Replicate Estimates Tables
(See page 5 about MOEs for Zero Counts and Percent Estimates of Zero or 100 Percent)

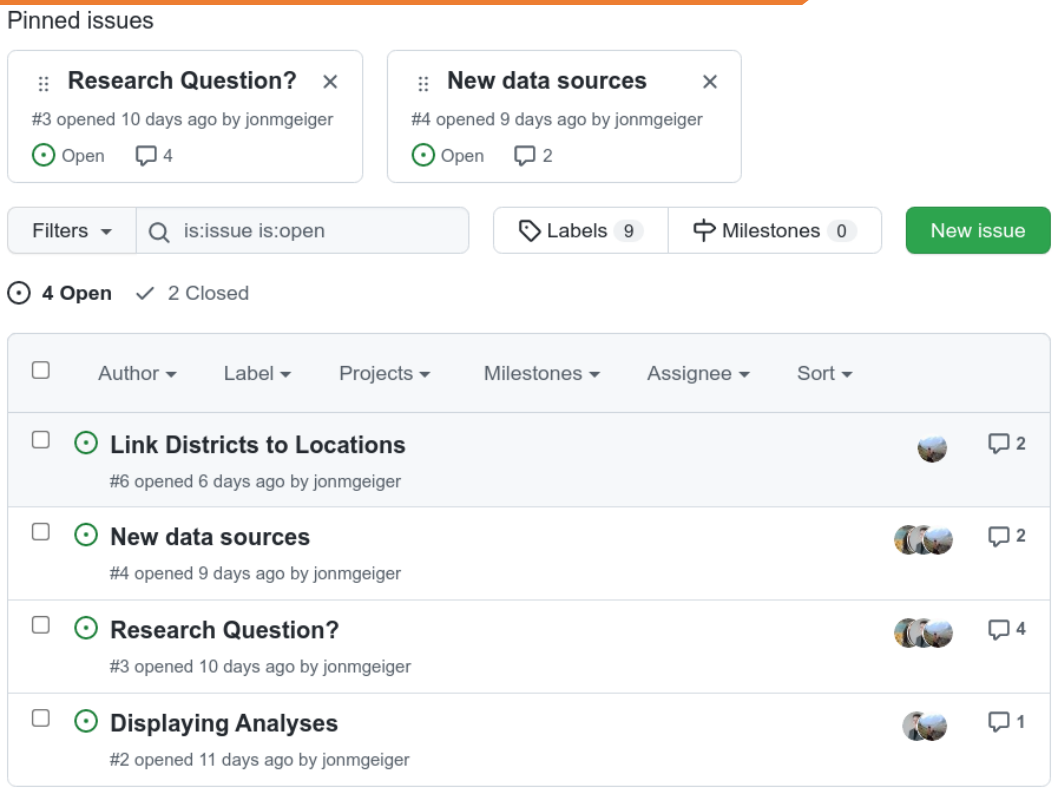
External Data Sets

In-progress Data Sets

- Racial/Ethnic Breakdown within School Districts
- Geographic Data
  - TIGER/Line files and Shapefiles
  - tigris R package

Organization/Management

GitHub!



Create an Issue → Assign people to that issue

People link relevant information → Create a branch and Pull Request linked to that Issue

Discuss changes in the Pull Request → Merge branch when finished

Working Research Questions

- What are some of the biggest influences on *graduation rates* of school districts in the U.S.?
- What are some of the biggest influences on *year-to-year retention rate* of school districts in the U.S.?
- To what extend did household conditions affect graduation rates before the spread of COVID compared to during the pandemic?
- How do household conditions compare to the funding each school receives, either independently or from the government?
- How much do racial ethnic demographics correlate with household conditions across school districts?

Difficulties & Next Steps

- Faulty Data - Sampling Bias?
  - Cutoff value for # children?
- Focused Research Question
  - What models should we be building?
- Including data in the repository
  - Working with TIGER/line files using tigris
- Displaying Analyses
  - HTML Files too large to view in GitHub natively

```
## $state
## [1] "New Mexico"
##
## $dist
## [1] "Española Municipal Schools"
##
## $children
## [1] 5860
##
## $pct_pov
## [1] 0.2536434
##
## $pct_SP
## [1] 0
##
## $SP_MOE
## [1] 0.5
##
## $pct_HHVJ
## [1] 0
##
## $HHVJ_MOE
## [1] 0.5
##
## $pct_CC
## [1] 0.02449738
##
## $CC_MOE
## [1] 0.005
```