

Census 2020— Counting Under Adversity



Constance F. Citro, *Senior Scholar*, CNSTAT

Senior Statisticians Society

October 27, 2020

A Perfect Storm for 2020

Pre-Corona Virus—

- **Citizenship question saga** > > *Erosion of public trust* (immigrant communities)

Data Collection During Corona Virus (MAF/self-response fine)—

- **Postponed nonresponse follow-up [NRFU]** (asked for *extension of deadlines* for apportionment/redistricting—ran into hurricane/wildfire season)
- **Administration interference** (political appointees, schedule, undocumented)
- **Back to original deadlines by compressing operations** (lawsuits)
- **Rushed post-collection processing** (duplicates? imputations?)

➤ Apportionment/Redistricting—

- **Lawsuits** (who counts? undocumented excluded? noncitizens excluded?)
- **Fitness for use** (how/when know quality? asterisked census? work-arounds?)
- **Confounding role of differential privacy protection?**

Census Bureau Future—

- **Need to restore credibility/trust?**

Scope of Talk

- Remind ourselves *why the census matters*
- Revisit *1920 census*
- Address *2020* and its *aftermath*
 - ❖ What expectations can we have about *data quality*?
 - ❖ What problems can we foresee for the *Census Bureau*?

Why the Census Matters

CONSTITUTIONAL USES

- Censuses not new idea: date to ancient Egypt to count people for taxation/military service (Latin, “censere,” to estimate); British carried out censuses in many American colonies
- US *first* with **CONSTITUTIONALLY MANDATED** census (Article 1, Section 2) to *peaceably* reapportion power as pop. grew and spread out (*seats/direct taxes/electoral college*)—*Everyone* counted (exc. tourists; Indians “not taxed”; slaves 3/5^{ths})
- **REDISTRICTING** *Congress, state legislatures, et al.* (1960s court cases, 1965 Voting Rights Act; 1975 law requires Census Bureau to provide small-area data to states)
- **2020**—Who counts? *Undocumented immigrants? Nor* (lawsuits)

2020—Which States May Gain/Lose Seats?

CHANGE IN SEATS	1	2	3
Gain	{Arizona	Texas	
West	{Colorado		
	{Montana		
	{Oregon		
	Florida		
	North Carolina		
Loss	{Illinois		
	{Michigan		
Rust Belt	{Minnesota		
	{New York		
	{Pennsylvania		
	{Rhode Island		
	{West Virginia		
	California		

Source: Election Data Services, December 2019

Why the Census Matters

KEY TO FEDERAL STATISTICS (*census/ postcensal estimates*)

- **Benchmarks** for surveys (ACS, CPS, NHIS, etc.)
- **Denominators** for rates (health, justice, etc.)
- **Components** of allocation formulas (billions \$
—Medicaid, education, highways)

KEY ISSUE—Will 2020 be good enough to rebase population estimates?

Why the Census Matters

CPS Male Coverage Ratios: March 2016 - March 2019



Why the Census Matters

KEY TRENDS (sense of who we are)

	1790	2019
Population	3.9 mil.	330 mil.*
House Seats	65	435 (since 1913)

*US third most populated country (after China, India)

	1850	1910	1970	2018-19
% People of color	16%	12%	16%	40%**
% Rural	85% (1790: 95%)	54%	26%	14%
% Foreign-Born	9.7%	14.7%	4.7%	13.7%

** Includes African-Americans, Asian-Americans, American Indians and Alaska Natives, Hispanics, and Native Hawaiians and Pacific Islanders

Why the Census Matters

Technological/Methodological Innovation—Census ever more **complex**

- 1890—processed using Hollerith **punch cards**
- 1940—**sampling** used to collect a few items (led to long-form sample; **ACS**)
- **1950**—(first commercial) **computer** used to process data (UNIVAC1)
- **1980**—capture/recapture wildlife estimation methods adapted for **coverage measurement** (dual-system estimation via post-enumeration survey— DSE/PES)

1920—An Earlier Storm

Aftermath of WWI/flu: spike in violence—riots, lynchings, labor strikes—fears of revolution

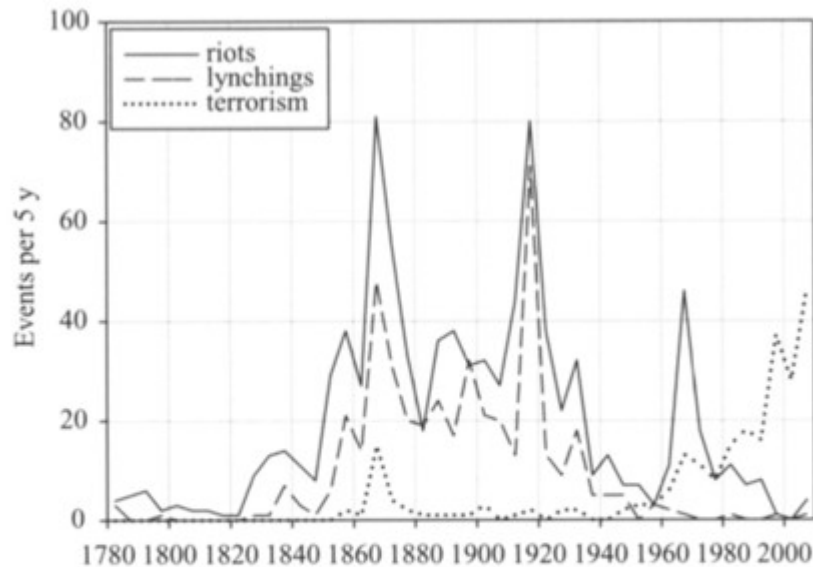


Figure 5. Dynamics of different types of political violence: riots, lynchings, and terrorism

The last category also include assassinations and rampages.

Peter Turchin. (2012). Dynamics of political instability in the United States, 1780-2010. *Journal of Peace Research* 49(4) pp. 577-591.

1920—An Earlier Storm

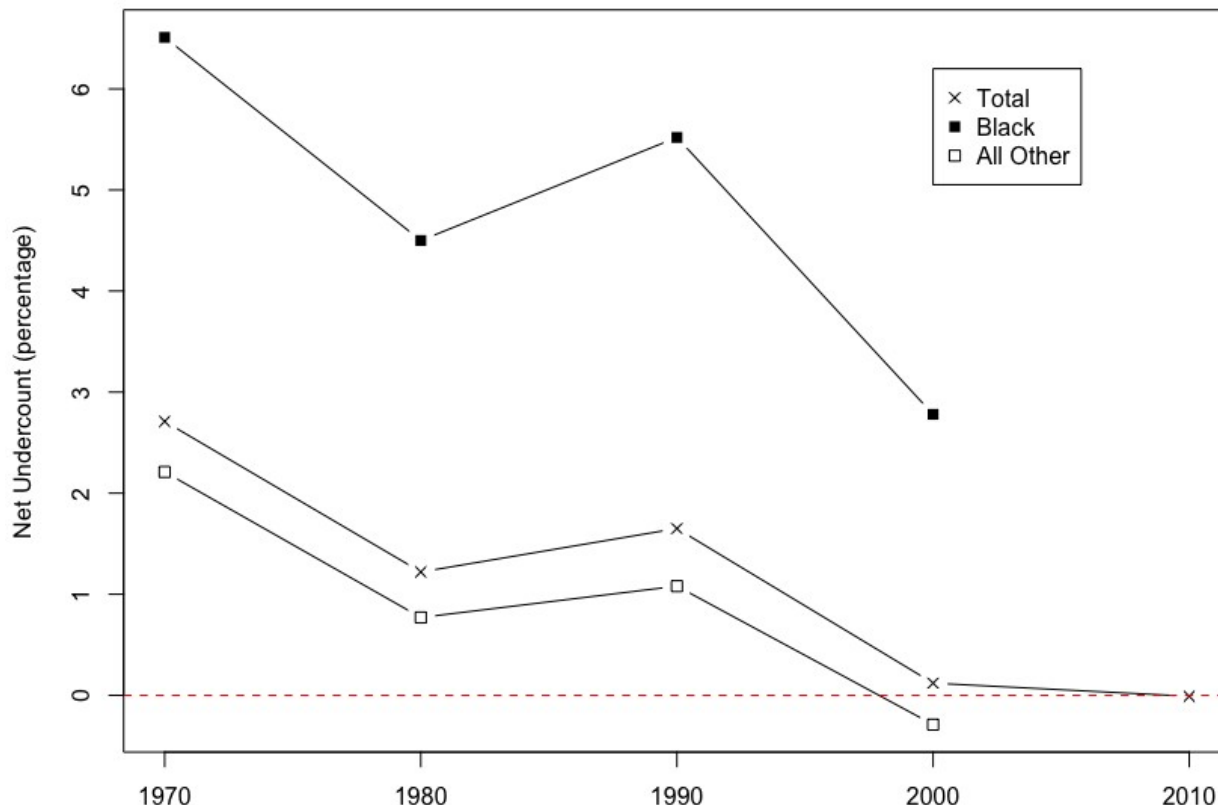
- *1920 Census* required *more time* (flu, winter weather, trouble hiring—took 5 > 1 mos.)
 - Rural states did not like results—*urban (majority); immigrants*
 - First time in U.S. history, Congress did **NOT** reapportion the House (fight about method, no more adding seats to cushion the pain)
 - Congress did pass **1924** law restricting immigration (made Native Americans citizens)
- Congress passed **1929** law providing for *automatic* reapportionment
 - Trade-off: States allowed to *malapportion* districts (1930s NYS districts ran from 799,407–90,671 pop.)
 - Redressed in 1960s *one-person/one-vote* court cases: made census block-level data *essential* for redistricting

1970-2010 ~ Net Undercount

Goal is count everyone *once* and *only once*, but always *duplicates*, *omissions*

Over time, net undercount *down*, but still *higher* for *people of color* than whites— what will *2020 rates* be??

Net Undercount Percentages, 1970-2010 Censuses



Measured by demographic analysis (DA); comparable estimates by race not available for 2010



2010~ Net Undercount

2010 Post-Enumeration Survey (PES)

$[(\text{PES minus census})/\text{PES}]$ (hh. population)

Total (-0.01%)

Overcounted groups		Undercounted Groups	
Non-Hispanic White	-0.84%	AI on Reservation	4.88%
Owner	-0.57	Non-Hispanic Black	2.07
Females 50+	-2.35	Hispanic	1.54
Update/Leave	-1.37	Renter	1.09
Midwest	-0.57	Kids 0-4	0.72
		Men 18-29	1.21
		Update/Enumerate	7.87

Estimates shown are statistically significantly different from zero (90%).

2020 Self-Response

➤ **2020** Internet response went smoothly (**1970** pioneered self-response via mail to improve quality/reduce costs)

➤ *Key for quality*

2010 mail responses (much better than enumerator returns)

- 98.1% correct enumerations
- 1.7% erroneous enumerations
- 0.2% whole-person imputations
- Mail returns also have lower rates of omissions than enumerator returns (based on 2000 analysis)

➤ *Final self-response rates* (% addresses)

2020 – 67.0% (53.5% via Internet); **2010** – 66.5%

Who Self-Responds? – 2010 Data

Total* 79.3%

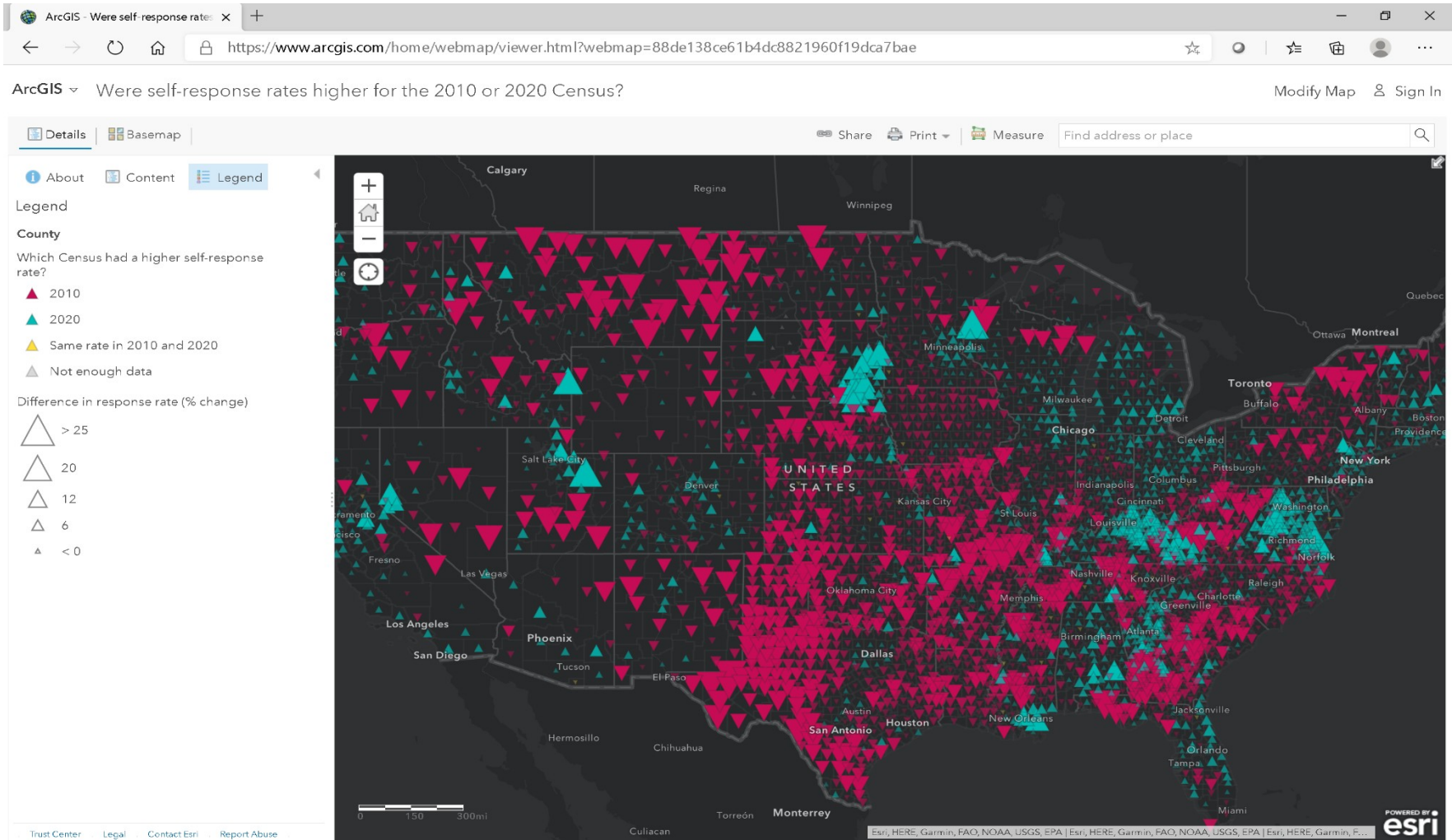
Highest (%)

Lowest (%)

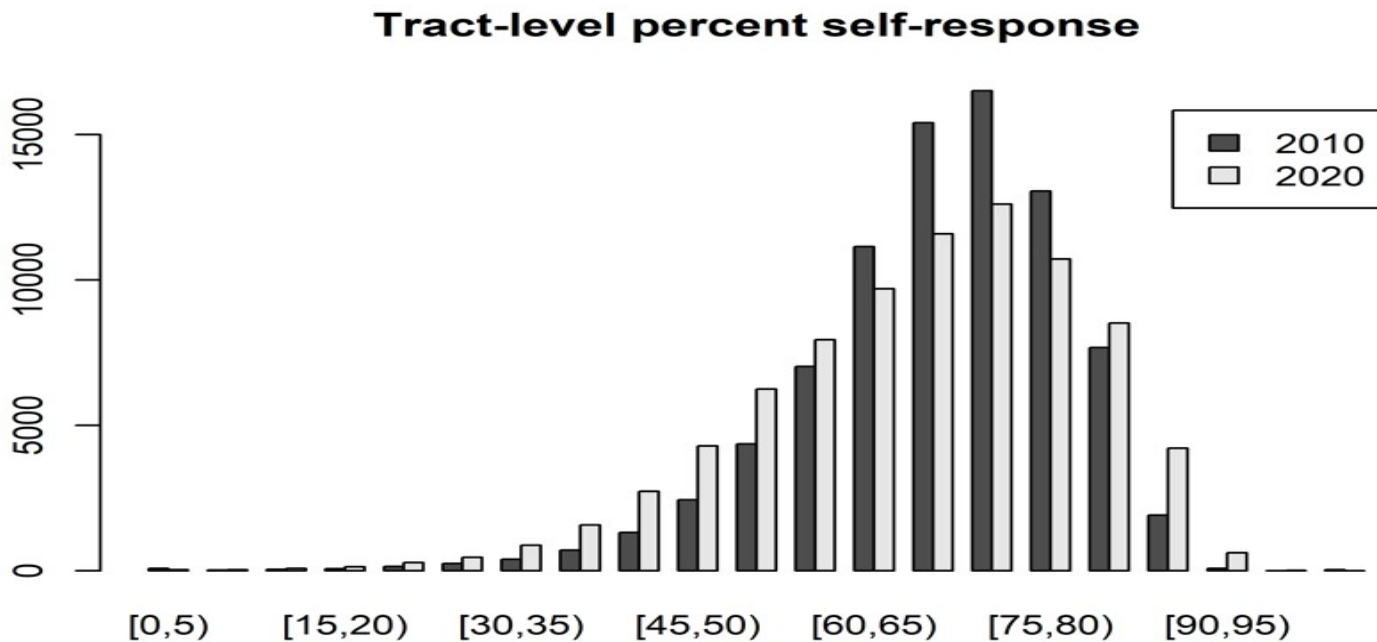
Age 65+	90.0%	18-24	55.4%
White	82.5	Hispanic	65.2
Owner	85.8	Renter	66.9
2 Person (hh.)	81.0	7+ Person (hh.)	67.0

*2010 mail response as percent of *occupied* households; self-response as percent of *all* addresses was 66.5%

2020 Self-Response (variable!)



2020 Self-Response (variable!)



Analysis by Bob Fay for ASA 2020 Census Quality Indicators Task Force

2020 (2010) Self-Response Rates

State		County	Census Tract	
NATION	67% (67) (2020—54% via Internet)			
AZ	64 (61)	Maricopa	68% (64)	
		Navajo*+	35 (40)	
		Apache*+	24 (40)	
DC	64 (66)		10.03	91%
			67.00	75
			88.03	41
MD	71 (70)	Montgomery	78 (76)	
		Baltimore City+	57 (60)	
		Worcester	33 (33)	
VA	71 (69)	Fairfax	81 (76)	
		Arlington	76 (73)	
		Lee+	41 (56)	

*Largely tribal lands +Total >13 pp Internet response

Enumeration Schedule(s)

Operation	Original End Date	New End Date**
Nonresponse Follow-Up (NRFU) (household or proxy response)	July 31, 2020	October 31, 2020
Group Quarters Enumeration	July 31, 2020	September 3, 2020
Visit Rural Households	April 30, 2020	July 29, 2020
Unduplication, Imputation, Review	December 31, 2020	April 30, 2021
* Deliver State Counts to President	December 31, 2020	April 30, 2021
Demographic Editing/Imputation, Privacy Protection, Tabulation	March 30, 2021	July 31, 2020
* Deliver Redistricting Data	March 30, 2021	July 31, 2021
Post-Enumeration Survey Results	June 30, 2021	November 30, 2021

* House Democrats okay with extending deadlines; nothing doing in the Senate; went back to original deadlines (or as close as possible).

** Dates shown announced in May 2020; Census (under pressure) sped up schedule to end NRFU by September 30; lawsuits; Supreme Court, October 13, said could stop NRFU; NRFU ended October 15.

Post-Collection Processing

- Takes *time*—Census wanted ~5 months; now <3
- Essential steps for apportionment (A)
 - *Unduplication* (non-IDs; >1 for same MAF ID; people enumerated in two locations—college students, snowbirds) (complicated by spring self-response/fall NRFU; hurricanes, wildfires)
 - *Count imputation* (for number people in household)
 - *Review* by professional staff (has to be short-circuited)
 - State counts may look okay
- Essential steps for redistricting (R)
 - *Whole-person/item imputation/editing* for characteristics (5% total race imputation in 2010—likely higher for 2020)
- *NEW*—Estimates for *undocumented* (A)/*voting age citizens* (R)
 - Alabama suing to limit (A) to *legal residents*; other states suing to *block* this (with Supremes now); MALDEF et al. suing to *block* CVAP estimates for (R)

Whole-Person Imputation Rates*

	2010		2000	
Count (household size NOT known; size/characteristics must be imputed)	1.2 mil.	0.4%	1.2 mil.	0.4%
Characteristics (hh. size known; age/race/ethnicity/sex must be imputed)	4.8 mil.	1.6	4.6 mil.	1.7
TOTAL	6.0 mil.	2.0	5.8	2.1

GROUPS with highest whole-person imputation rates (2010)

AIAN on Reservations	4.1%
African-American, Ages 18-29, NHPI, Other Race, Renter	2.8-3.1%

*Disproportionately from proxy enumerations (landlord, neighbor)

2020 Data Quality??

➤ *Demographic analysis (DA)*—

- Out in mid-December 2020, but can't tell us much

➤ *PES/DSE*—

- Needed for net undercount, duplications, omissions
- No results until late 2021, quality may be iffy

➤ *Process indicators*—

- Whole-person imputations, proxy enumerations, duplicates eliminated, etc., etc.
- Can/will they be provided expeditiously for small areas/population groups? (typically not available for a couple of years, little detail)
- Comparisons needed with 2010

ASA Recommendations for Quality Indicators

- We recommend a series of *quality indicators* to be used to assess the 2020 Census. . . .
- The indicators should be *readily available* and *used expeditiously* to assess the quality of the 2020 Census.
- Qualified *external researchers* should be granted access to the data to help conduct the analyses.
- *Additional assessments* should be conducted when more data become available.
- Early planning for the *2030 Census* should build on the lessons of 2020, be conducted in public, and include extensive stakeholder input.
- The Census Bureau's authorizing legislation, *Title 13 U.S. Code*, should be updated.
 - First, Title 13 needs to align better with recent *Information Quality Act and OMB guidelines*.
 - Secondly, Title 13 should also be examined to determine if *further amendments to protect the integrity and independence* of the Census Bureau and the decennial census are warranted.

My Concerns/Questions (partial list!)

- *Transparency re quality*—can Census speed up release of process indicators for small areas/groups?
- *Differential privacy protection*—will DP impair ability to use the data/understand data quality?
- *Pop estimates*—will/can Census use alternate data sources for pop estimates over the decade?
- *Undocumented immigrants*—how can/should the Census Bureau deliver numbers?
- *Legal structure*—how can/should laws be beefed up to protect Census Bureau/census data from political malpractice?

Look Forward to Discussion

Connie Citro, Senior Scholar, CNSTAT
ccitro@nas.edu

**CNSTAT reports on census methodology
available from the National Academies Press**
<http://www.nap.edu>



References

- American Statistical Association. (October 13, 2020). *2020 Census Quality Indicators: A Report from the American Statistical Association*. Prepared by the 2020 Census Quality Indicators Task Force.
<https://www.amstat.org/asa/files/pdfs/POL-2020CensusQualityIndicators.pdf#:~:text=The%20Census%20Quality%20Indicators%20Task%20Force%20%28task%20force%29,advocating%20for%20and%20contributing%20to%20an%20accurate%20census>
- ARC GIS (ESRI). (2020). *Were self-response rates higher for the 2010 or 2020 Census?* (data by state, county, census tract).
<https://www.arcgis.com/home/webmap/viewer.html?webmap=88de138ce61b4dc8821960f19dca7bae>
- Keller, Andrew, and Fox, Tyler. (May 22, 2012). *2010 Census Coverage Measurement Estimation Report: Components of Census Coverage for the Household Population in the United States*, U.S. Census Bureau.
<https://www2.census.gov/programs-surveys/decennial/2010/technical-documentation/methodology/g-series/g04.pdf>
- Mule, Thomas. (May 22, 2012). *2010 Census Coverage Measurement Estimation Report: Summary of Estimates of Coverage for Persons in the United States*. U.S. Census Bureau.
<https://www2.census.gov/programs-surveys/decennial/2010/technical-documentation/methodology/g-series/g01.pdf>
- NPR, Hansi Lo Wang (writes/tweets on the 2020 census).
<https://www.npr.org/people/177498291/hansi-lo-wang/archive>