Datasets

Census Program Pages

The home pages for each census program. These contain news, release schedules, methodology, technical documentation, and tutorials. Direct downloads for smaller datasets are also available from these pages.

Decennial Census
https://www.census.gov/programs-surveys/decennial-census.html

American Community Survey
https://www.census.gov/programs-surveys/acs/

Population Estimates
https://www.census.gov/programs-surveys/popest.html

Business Patterns
https://www.census.gov/programs-surveys/cbp.html

Economic Census
https://www.census.gov/programs-surveys/economic-census.html

Decennial Census

The DEC is the cornerstone of the Census Bureau's programs. It is the original census dataset that has been collected since 1790, as stipulated by Article 1, Section 2 of the Constitution for the purpose of reapportioning seats in Congress between the states and redistricting Congressional districts within states. The DEC also serves as an input and a control for many other census data programs, such as the ACS and Population Estimates, as well as for datasets generated by other federal agencies.

The DEC is a 100 percent count of the entire population, conducted every decade in years ending in zero. It is conducted over a period of several months but uses a reference date of April 1 for determining residency. The concept of usual residence is employed, where each household is asked whether their current address is the place they usually live on April 1 (note that this is quite different from asking what a person's permanent or legal residence is). With the exception of foreign tourists and visitors on short-term visas, every person who resides in the United States is counted regardless of their status or living circumstances. The Census Bureau maintains a master address file for the entire nation, and all households are sent a census form that they are required by law to submit. For the 2020 census, people were able to submit their form online for the first time. The Census Bureau works with administrators of group-quarters facilities to count populations in these facilities and has a dedicated program for counting the homeless population.

The Bureau sends enumerators to certain hard-to-count areas to collect data in person and to households or quarters that don't respond to the initial form or follow-ups. The counting process for 2020 was noteworthy as it coincided with the COVID-19 pandemic, social unrest, and several severe weather events. The self-response rate for the 2020 census was 67 percent, which was comparable to the 2010 census, and after nonresponse follow-up, the Bureau estimates that over 99 percent of all addresses were reached (US Census Bureau 2020a). The count was noteworthy in that approximately 80 percent of all self-responses were submitted online and because item nonresponse (where a household left certain questions blank) was higher than in the past (US Census Bureau 2021d). While data for the vast majority of Americans is captured by self-response and nonresponse follow-up, the Bureau applies a number of imputation techniques to estimate missing data for those who are not directly reached.

Also new for 2020 is the adoption of a differential privacy mechanism, a mathematical approach with roots in cryptography that adds slight alterations or "noise" to published summary data and microdata to prevent reverse engineering of the census (in concert with third-party datasets) to reveal information about individuals. While the Census Bureau emphasizes that these new techniques are necessary given the current data privacy landscape and that the impact on most analyses is quantifiably minimal (US Census Bureau 2021b, 2021c), some demographic experts and policy makers have been critical of the changes (Wang 2021; IPUMS 2021). Another aspect of the new mechanism is a reduction in the number of tables that will be published and a decrease in their level of geographic detail, with fewer tables available at the block, block group, and tract levels.

Once the count is finished and the data is published, the Bureau uses two different techniques for estimating a possible undercount or overcount: a post-enumeration survey to a sample of households and demographic analysis using birth, death, and migration data from the Population Estimates program. Relatively speaking, the DEC is highly accurate; both the 2000 and 2010 census slightly overcounted the total population, while the 2020 census neither undercounted nor overcounted the total. However, in each case there was a differential undercount, where certain groups were missed while others were double-counted; young children, minorities, and people of low income tend to be undercounted.

The data is released in stages. By law the official population count of the US and each state must be submitted to the president by December 31 of the census year to launch the reapportionment process. The following March, the Public Redistricting files that contain basic population and housing counts down to the block level are released so that the redistricting process can begin. The following summer and fall, the most detailed census data is released on a rolling basis as part of the Summary File 1 series, with additional cross-tabulations by race and ethnicity published in Summary File 2 the following year (for the 2020 census, the term *Summary File* is slated to be replaced by the term *Demographic and Housing Characteristics File*). The timetable for the release of all the 2020 census products was delayed due to the COVID-19 pandemic and political upheaval. Apportionment data wasn't released until March 2021, and redistricting data was finally published in August 2021. The data shows that the population grew from 309 to 331 million between 2010 and 2020. This increase of 7.4 percent was the lowest rate of growth since the 1930s.

The contemporary DEC dataset consists of just basic demographic characteristics collected through the 100 percent count. These variables include age, sex, race, ethnicity, household and family relationships, group-quarters population, and housing unit occupancy and tenure. This focus on the basics was introduced in the 2010 census and continues for 2020. In contrast, from the 2000 census backward there were two questionnaires: a short form that collected the basics from the entire population and a long form that was sent to a large sample of the population (one in six households) that collected detailed socioeconomic and housing characteristics such as educational attainment, employment status, home value, and much more. The ACS was introduced in 2005 to capture all of these detailed characteristics on an annual basis so that in 2010 the decennial census could revert to being a simpler count.

Researchers should use the decennial census when they

- need precise counts of the population
- need block- and block-group-level data
- require just basic demographic variables
- do not require current data (particularly by decade's end)
- are studying small population groups
- are doing historical comparisons with prior decennial censuses

The DEC will not be your choice if you need detailed socioeconomic or housing characteristics, current data (particularly toward the end of a decade), or annual data.

American Community Survey

The ACS was launched in 2005 to provide detailed population and housing characteristics on an ongoing basis. It was designed to provide data in a way that was more timely and cost-effective than the DEC, while allowing the latter to focus more exclusively on counting the basic characteristics of the population. The ACS is a rolling sample survey of 292,000 addresses a month, for a total of 3.5 million addresses per year. The sample is stratified so that a certain number of addresses in every state are captured. The concept of current residence is used for determining residency, where a household is asked whether they are staying at the current address for at least two months. Sampled households are required by law to return the survey by mail or online, and the sample includes about 2.5 percent of all group-quarters addresses in the Bureau's master address file. Unlike the DEC, the Bureau follows up with only a sample of nonresponders and uses imputation techniques only for incomplete forms.

Each year, the sample data collected for the previous year is used to generate estimates for the total

Table 4.1. ACS Table Prefix Codes and Subjects

ID	Subject	ID	Subject
00	Unweighted Count (of the Sample)	15	Educational Attainment
01	Age and Sex	16	Language Spoken at Home
02	Race	17	Poverty Status
03	Hispanic or Latino Origin	18	Disability Status
04	Ancestry	19	Income
05	Citizenship Status, Year of Entry, Foreign Born	20	Earnings
06	Place of Birth	21	Veteran Status
07	Migration and Residence 1 Year Ago	22	Food Stamps/SNAP
08	Commuting and Place of Work	23	Employment and Work Status
09	Relationship to Householder	24	Industry, Occupation, Class of Worker
10	Grandparents and Grandchildren	25	Housing Characteristics
11	Household and Family Type	26	Group Quarters
12	Marital Status and History	27	Health Insurance Coverage
13	Fertility	28	Computer and Internet Use
14	School Enrollment	29	Citizen Voting-Age Population

population. A one-year period estimate is published for all geographic areas of the country that have at least 65,000 people. As the sample size is too small to publish reliable annual estimates for smaller areas, sample data for five years is aggregated to produce a five-year period estimate for all geographic areas down to the block group level. Each year, a new five-year period estimate is released by dropping out samples from the oldest year and adding samples from the latest year.

ACS statistics are published as estimates with margins of error at a 90 percent confidence level. For example, according to the 2019 ACS, the population of Providence, Rhode Island, was 179,875 +/− 40. This means that we are 90 percent confident that the population is somewhere between 179,835 and 179,915 people, and there is a 10 percent chance that the actual population falls outside this range. The smaller the population group or geographic area, the higher the margin of error will be. For instance, the population under age eighteen in Providence in 2019 was 36,020 +/− 3,401. In all published tables, ACS estimates are always provided in pairs with the estimate in one column and the margin of error in a second column. Estimates representing percent totals are also published with a margin of error.

Unlike the DEC statistics, ACS estimates are fuzzy intervals that can be used for generally characterizing an area and should never be interpreted as exact counts (Spielman, Folch, and Nagle 2014). This fuzziness is even greater for areas with less than 65,000 people, as we can characterize only a five-year time frame. For example, estimates for Newport, Rhode Island, are published for only a five-year period because the population of the city falls below the 65,000 threshold. We can say that the population of the city was approximately 24,663 +/− 34 between the years 2015 and 2019. If we wanted to make comparisons between Newport and Providence, we would need to use data from the five-year ACS for both places, since data is unavailable for Newport in the one-year series.

The great strengths of the ACS are that it provides annually updated data for large and small areas of the US and includes a broad range of detailed demographic and socioeconomic characteristics that are not captured in other census or government datasets. Table 4.1 summarizes the ACS subject categories and their prefix codes, which are used to identify and group the tables.

Its great weaknesses are its complexity, which makes it more challenging to work with, and reliability issues that are serious for small areas and population groups. It is not uncommon that ACS data is misinterpreted and used incorrectly. In particular, the margin of error is often ignored or misunderstood, and the data is inappropriately treated as a count (Jurjevich et al. 2018; Nesse and Rahe 2015). Ignoring the margin of error can lead to faulty conclusions regarding the significance of findings (Jung, Thill, and Issel 2019). Even though 3.5 million addresses are sampled, due to nonresponse and other issues, just over 2 million addresses are used to generate estimates in any given year. The level of uncertainty in the ACS has exceeded the Census Bureau's initial expectations, and some researchers have suggested that the block-group-level data is unusable for most applications, and even census-tract-level data does not provide reasonable certainty in estimating certain characteristics (Spielman, Folch, and Nagle 2014; Salvo et al. 2007). Due to a high nonresponse rate caused by disruptions from the COVID-19 pandemic, the Census Bureau is not publishing traditional one-year estimates for the 2020 ACS but is releasing a smaller number of experimental tables for large geographic areas instead (US Census Bureau 2021a).

Researchers can make compromises to counter the effect of low precision, essentially by using estimates generated from a larger sample. One approach is to use a five-year estimate even though a one-year estimate for a given place is available. For example,

the one-year 2019 estimate for the population under eighteen in Providence was 36,020 +/− 3,401, but if we used the five-year 2015–2019 estimate, the under-eighteen population was 40,156 +/− 1,006, a much smaller margin of error relative to the estimate. Another approach would be to use a smaller number of subcategories for a given variable, such as households by income classified into fewer income brackets. The Census Bureau provides some of these aggregates in the collapsed C tables (described in the previous chapter) for that purpose.

A third option would be to use a larger geography, sacrificing geographic detail for better estimate precision. Users can aggregate geographic areas or subject categories and use formulas to calculate margins of error for new estimates. For example, to add the population for two geographies together, you would sum the estimates and then take the square root of the sum of the squares of their margins of error to calculate a margin of error for the new estimate. For example, the Society Hill neighborhood in Philadelphia falls largely within census tracts 10.01 and 10.02. Using the 2015–2019 ACS, to create an aggregated neighborhood estimate of the number of people who take public transit to work, you would sum this population from Table B08006 for the two tracts:

```
273 + 436 = 709
```

And then take the square root of the sum of the squares of their margins of error:

```
√((80^2)+(151^2)) = 171
```

A statistic called the coefficient of variation (CV) is used to characterize the reliability of an estimate. To calculate a CV, you divide the margin of error by 1.645 (the constant *Z* value for a 90 percent confidence level), divide that result by the estimate, and then multiply by 100. Opinions vary on what an acceptable CV level is; as a general rule of thumb, CVs of 0 to 12 can be considered as highly reliable, 13 to 34 as medium reliability, and 35 and above as low reliability.

To calculate the CV for public transit commuters in the Society Hill neighborhood:

```
((171/1.645)/709)*100 = 15
```

This estimate would be interpreted as having medium reliability, as it has a value between 13 and 34. The CVs for the individual tracts, 10.01 and 10.02, are 18 and 21, respectively. While these are also of medium reliability, the aggregated neighborhood CV is lower and thus is considered to be more reliable.

A number of tools and resources exist to help researchers understand and work with ACS estimates (see the tools mentioned in the gray box under ACS Resources), and the Census Bureau provides recommendations with detailed examples in a series of guidebooks (US Census Bureau 2020b).

It's particularly important to recognize the fuzziness of estimates when comparing data over time. Given that many differences in values from one year to the next are more likely due to sampling variability rather than actual change, the one-year estimates do not lend themselves to being studied as a time series. There is a test for statistical difference that allows you to determine whether two different estimates are truly different. For the five-year estimates, only nonoverlapping periods should be compared; otherwise you would be comparing estimates that are generated from the same sample pool. For example, comparing estimates from 2010–2014 and 2015–2019 would be sound, as there is no sample overlap between the two series.

Researchers should use the ACS when they

- need detailed socioeconomic characteristics of the population
- are doing historical comparisons of these characteristics that are not captured in recent decennial censuses
- need current data
- can live with the fuzziness of the estimates

The ACS will not be your first choice if you need precise counts (particularly for small population groups and geographies), if you are studying basic demographic variables that are available in the DEC and timeliness is not an issue, or if you want an annual time series.

ACS Resources

ACS Calculator
Online calculators for significant difference and computing a new estimate from two existing estimates, from the Cornell Program on Applied Demographics
https://pad.human.cornell.edu/acscalc/index.cfm

ACS Tools
Excel spreadsheets with pre-built macros for significant difference and calculating new estimates, from the Economic, Demographic, and Statistical Research Unit of Fairfax County, VA (scroll to the bottom for the tools)
https://www.fairfaxcounty.gov/demographics/research-tools

Population Estimates

Whereas the DEC is a count and the ACS is a sample survey, the Population Estimates are generated from administrative records using a series of calculations. Using the DEC as a base, components (births, deaths, and migration) for different cohorts of the population (age, sex, race, Hispanic origin) are used to create new estimates for successive years. The cohort-component method is a basic approach used in demography; for a given place, data on births and deaths is used to calculate how many people were born in the first age group and how many people survived and thus moved from their current age group to the following one. Different sources for domestic and foreign migration data are used to determine how many people moved into and out of an area. The end result is a new estimate of the population for that given year. The process is repeated the following year using the previous year as a base.

The Census Bureau uses this approach to generate estimates for the nation, states, counties, and metropolitan areas each year. It is an iterative process carried out in stages to ensure that the sum of smaller geographies equals the whole. The data is produced in two sets: estimates of the components (births, deaths, migration) and of the characteristics (age, sex, race, Hispanic origin). A different methodology is used to generate basic population totals for places (cities and towns) and an estimate of total housing units for all geographies.

The Census Bureau releases the estimates in a series called a vintage. Each vintage contains the latest year of estimates, plus all previous years back to the previous DEC. The estimates for previous years may differ from previous vintages, as the Census Bureau modifies and improves its models and methodology based on new information. This means that an estimate for the year 2018 that appears in the 2019 vintage may be different from the estimate for 2018 that appeared in the 2018 vintage. For that reason, it's important for users of this data to wholly replace each vintage with the subsequent one. At the beginning of a new decade, a final vintage based on the DEC from ten years past is used to gauge the accuracy of the latest DEC. Following this release, two things happen. First, using the latest DEC the Bureau will go back and revise the annual estimates for the previous decade for a final time, based on what the latest count shows. This series is referred to as the Intercensal Estimates, and it becomes the definitive estimate series for that decade. Second, the subsequent vintage estimates for the next decade use the new DEC as the estimates' base.

Compared to the DEC and ACS, the estimates are a much smaller data series with far fewer tables and variables, and thus the dataset is easier to access and work with. In addition to data.census.gov and the API, users can easily search and browse through spreadsheets published on the Population Estimates Program website. The series is well suited for studying annual population change and basic demographic characteristics for large areas. The estimates are the only series that contains the components of population change—births, deaths, net domestic migration, and net foreign migration—and thus are the source for researchers interested in studying these components. Natality and mortality data in the US is compiled at the county level by local health departments, is reported up to the state level, and is collated at the federal level. Domestic migration data is sourced from the IRS and the Social Security Administration, while foreign migration data comes from multiple sources based on migration origin (i.e., US military overseas, US territories, foreign countries).

Researchers should use the Population Estimates when they

- need basic population counts and characteristics for large areas
- are studying annual population change
- need data on the components of population change

The population estimates will not be your first choice if you need more detailed characteristics or data for small geographic areas.

Business Establishments

The Census Bureau has been collecting and summarizing data on businesses almost as long as it has been collecting population data (Micarelli 1998). There are several different statistical programs that measure business activity, and in this section we will discuss two of the main programs: the Business Patterns and the Economic Census.

Both of these programs compile statistics on business establishments, which are defined as single physical locations where business is conducted or where services or industrial operations are performed. Establishments are assigned to industries, which are groups of businesses that produce similar products or provide similar services, using the North American Industrial Classification System (NAICS). Introduced in 1997, NAICS is used by government agencies in the US, Canada, and Mexico for classifying industrial activity, replacing the earlier Standard Industrial Classification (SIC) system. NAICS assigns businesses into broad groups and detailed divisions and subdivisions, with two-to-six-digit codes that indicate related groupings and the level of detail. Table 4.2 illustrates the different levels of NAICS for the Offices of Physicians Industry, with data from the 2019 Business Patterns for Rhode Island. This industry includes

Table 4.2. NAICS Codes illustrated for the Offices of Physicians Industry in Rhode Island, 2019 Business Patterns

NAICS Code, Title, and Level	Establishments	Employment
62 Healthcare and Social Assistance [sector]	3,113	87,067
- 621 Ambulatory Healthcare Services [subsector]	1,978	27,680
-- 6211 Offices of Physicians [industry group]	648	8,953
--- 62111 Offices of Physicians [industry]	648	8,953
---- 621111 Offices of Phys. (except mental health) [U.S. industry]	607	8,786
---- 621112 Offices of Phys., mental health [U.S. industry]	41	167

establishments where doctors are primarily engaged in the independent practice of general or specialized medicine or surgery, operating in private or group practices in their own offices or in the facilities of others, such as hospitals or medical centers.

The NAICS is revised every five years, in years that coincide with the Economic Census, to adapt the categories to fit a changing economy. Throughout NAICS there are miscellaneous categories that capture business activities that are not explicitly captured elsewhere. If an industry grows in importance over time, it may emerge as a distinct category in a revision, whereas an industry that declines in importance may be folded into a miscellaneous group. Each version of the NAICS is named for the year of its revision (e.g., NAICS 2017 and NAICS 2022). Concordances or crosswalks are published that allow researchers to relate the changing categories over time. The Economic Census and the Business Patterns use the NAICS version that coincides with their years (i.e., the 2017 Economic Census and the 2017 to 2021 Business Patterns use NAICS 2017).

The Business Patterns is an administrative dataset generated annually from the Business Register, a federal database of all businesses with paid employees in the United States. It is often referred to as the County Business Patterns and ZIP Code Business Patterns, the two most commonly used geographies in the series (which also includes states, metro areas, and Congressional districts). Unlike the population datasets, the Business Patterns and Economic Census report data by actual USPS ZIP codes instead of ZCTAs, as the ZIP information is simply scraped from the address record of the business. The Business Patterns series is relatively small and includes summary data on employees, establishments, and payroll by NAICS and geographic area. There are also tables that count establishments by employee size and as multiunit firms.

The Economic Census is a larger undertaking that generates data from counts, sample surveys, and administrative records. It is conducted every five years, in years ending in two and seven. It captures the same indicators included in the Business Patterns, as well as data on production and sales. Data is also published for the same geographies, plus cities and towns, and there are industrial as well as geographic summaries. The Economic Census is often used for studying trends as opposed to providing current statistics, as it takes several years from the time data is compiled to the time it is released. There are some slight differences between what industries the Business Patterns and Economic Census cover. Neither captures agriculture, government, or independent consultants and contractors, which are covered in different data programs (the US Department of Agriculture's Census of Agriculture and the US Census Bureau's Census of Governments and Nonemployer Statistics).

NAICS Reference and Other Business Datasets

NAICS
https://www.census.gov/naics/

Census of Agriculture
https://www.nass.usda.gov/AgCensus/

Census of Governments
https://www.census.gov/programs-surveys/cog.html

Nonemployer Statistics
https://www.census.gov/programs-surveys/nonemployer-statistics.html

Bureau of Labor Statistics Data Portal
https://www.bls.gov/data/

Like the population and housing datasets, the business datasets also fall under confidentiality regulations. Data on individual businesses is never reported; all data is summarized by geography and NAICS. There are a number of disclosure mechanisms with varying degrees of restriction designed to prevent users from reverse engineering the data to obtain establishment-level information. The Census Bureau may choose to inject noise into the employment or payroll values by inflating or deflating them by up to 10 percent, or may publish a range of values instead of a specific one, or may not publish the data at all if there are fewer than three establishments in a given industry or area or if one establishment comprises a large majority of employment or payroll. These policies present a range of challenges to data users. The disclosure mechanisms are indicated with footnotes that are embedded

in the data in place of actual values, which can create data processing challenges. Also, the sum of smaller parts will seldom equal the whole; if data is not disclosed at a six-digit NAICS level, it may be included in the four-digit level above it, or if data for an industry is not disclosed for a county, those statistics may be included in the state totals. Thus, aggregating data from smaller to larger parts leads to omission errors.

Both of these datasets count establishments and employees based on the location of the establishment. Beyond these series, there are other census datasets, like the ACS and CPS, that measure the labor force and typically count workers where they live (and in a few cases based on where they work). The Bureau of Labor Statistics, which collaborates with the Census Bureau in producing the CPS, also has a number of different data and survey programs for measuring business and labor activity.

Researchers should use the Business Patterns when they

- want annual data on establishments, employment, and wages by industry for large geographies
- want business data for ZIP codes and Congressional districts not published elsewhere

Researchers should use the Economic Census when they

- are doing long-term research, where timeliness is less important
- need data on production and sales for certain industries
- need data for ZIP codes, towns, and cities not collected elsewhere

References

IPUMS. 2021. "Changes to Census Bureau Data Products." IPUMS website. Minneapolis: Minnesota Population Center. https://www.ipums.org/changes-to-census-bureau-data-products.

Jung, Paul H., Jean-Claude Thill, and Michele Issel. 2019. "Spatial Autocorrelation and Data Uncertainty in the American Community Survey: A Critique." *International Journal of Geographical Information Science* 33, no. 6: 1155–75.

Jurjevich, Jason R., Amy L. Griffin, Seth E. Spielman, David C. Folch, Meg Merrick, and Nicholas N. Nagle. 2018. "Navigating Statistical Uncertainty: How Urban and Regional Planners Understand and Work with American Community Survey (ACS) Data for Guiding Policy." *Journal of the American Planning Association* 84, no. 2: 112–26.

Micarelli, William F. 1998. "Evolution of the United States Economic Censuses: The Nineteenth and Twentieth Centuries." *Government Information Quarterly* 15, no. 3: 335–77.

Nesse, Katherine, and Mallory L. Rahe. 2015. "Conflicts in the Use of the ACS by Federal Agencies." *Population Research and Policy Review* 34: 461–80.

Salvo, Joseph J., Arun Peter Lobo, Adam L. Willett, and Joel A. Alvarez. 2007. *An Evaluation of the Quality and Utility of ACS Five Year Estimates for Bronx Census Tracts and Neighborhoods*. New York: Population Division, New York City Department of City Planning. https://www1.nyc.gov/assets/planning/download/pdf/data-maps/nyc-population/eval_quality_utility_acs_five_year_estimates_bx.pdf.

Spielman, Seth E., David Foch, and Nicholas N. Nagle. 2014. "Patterns and Causes of Uncertainty in the American Community Survey." *Applied Geography* 46: 147–57.

US Census Bureau. 2020a. "2020 Census Response Rate Update: 99.98% Complete Nationwide." Press release no. CB20-CN.120, October 19. https://www.census.gov/newsroom/press-releases/2020/2020-census-all-states-top-99-percent.html.

———. 2020b. *Understanding and Using American Community Survey Data: What All Data Users Need to Know*. Washington, DC: US Department of Commerce, September. https://www.census.gov/programs-surveys/acs/guidance/handbooks/general.html.

———. 2021a. "Census Bureau Announces Changes for 2020 American Community Survey 1-Year Estimates." Press release no. CB21-123, July 29. https://www.census.gov/newsroom/press-releases/2021/changes-2020-acs-1-year.html.

———. 2021b. *Comparing Differential Privacy with Older Disclosure Avoidance Methods*. D-FS-GP_EN-0509. Washington, DC: US Department of Commerce, August. https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance.html.

———. 2021c. *Protecting the Confidentiality of the 2020 Census Redistricting Data*. D-FS-GP-EN-049. Washington, DC: US Department of Commerce, August. https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance.html.

———. 2021d. "2020 Census Operational Quality Metrics: Item Nonresponse Rates." *Random Samplings* (blog), August 25. https://www.census.gov/newsroom/blogs/random-samplings/2021/08/2020-census-operational-quality-metrics-item-nonresponse-rates.html.

Wang, Hansi Lo. 2021. "For the U.S. Census, Keeping Your Data Anonymous and Useful Is a Tricky Balance." NPR, August 2. https://www.npr.org/2021/05/19/993247101/for-the-u-s-census-keeping-your-data-anonymous-and-useful-is-a-tricky-balance.