Journal of Data Science 21 (2), 239–254
April 2023
DOI: 10.6339/22-JDS1082
Data Science in Action

Creating a Census County Assessment Tool for Visualizing Census Data

Isabel Youngs[1,*], Ronald Prevost[1], and Christopher Dick[1]
[1]*Georgetown University, 3700 O St NW, Washington, DC, United States*

Abstract

The 2020 Census County Assessment Tool was developed to assist decennial census data users in identifying deviations between expected census counts and the released counts across population and housing indicators. The tool also offers contextual data for each county on factors which could have contributed to census collection issues, such as self-response rates and COVID-19 infection rates. The tool compiles this information into a downloadable report and includes additional local data sources relevant to the data collection process and experts to seek more assistance.

Keywords *census; data visualization; data quality; tableau*

1 Introduction

Accurate decennial census data are critical to making effective decisions at the national, state, and local levels. The U.S. Census Bureau's population counts effect congressional redistricting, social program funding, school planning, and more. The importance of the decennial census is clear, but the 2020 Census faced unprecedented challenges with unknown impacts on data quality, from the COVID-19 pandemic and related stay-at-home orders to new household compositions, displacement of college students, natural disasters, and political interference. These factors led data users to question whether the data are fit to use.

Furthermore, the Census Bureau implemented a new framework of data privacy for the 2020 Census called "differential privacy". Changes to the Disclosure Avoidance System (the system by which the Census Bureau ensures that data releases protect respondent privacy) had a negative impact on data quality for less populated areas. Despite these challenges to data quality, there exist very few resources to understand which sub-state geographies and regions may have suffered from these impacts.

Data quality can be assessed in two major ways: by evaluating operational metrics or comparing the population count to other population totals. However, the Census Bureau has only released a handful of operational metrics, and they are limited to state level metrics. For population comparisons, the U.S. Census Bureau released the results of a Post-Enumeration Survey (PES) in 2022 (US Census Bureau, 2022). The PES is a survey given to a sample of the population after the decennial census to estimate the error. The most granular geographic results from this survey are also released at the state level, which is not detailed enough for many state and local data users. Another assessment method the Census Bureau developed to measure the quality of the decennial census is Demographic Analysis. Demographic Analysis uses

*Corresponding Author. Email: iwy2@georgetown.edu.

Received July 28, 2022; Accepted December 14, 2022

administrative data to produce national estimates of the population, which can be compared to the actual count to estimate coverage error. Unfortunately, Demographic Analysis also does not offer sub-state metrics of census data quality.

To assist state and local groups in assessing sub-state data quality indicators, we created the 2020 Census County Assessment Tool (CCAT). Similar to Demographic Analysis, we compared the actual census count to population estimates for total, group quarters, race and ethnicity characteristics, as well as housing estimates produced annually by the Census Bureau at the county level. The CCAT was designed to visualize whether the 2020 Census diverged significantly from the 2020 estimates. Additionally, our team compiled data from public and administrative sources to provide context for these differences in population counts and population estimates. Overall, while no single analysis can provide a true look at the complete quality of the 2020 Census, our goal in this project was to provide data users with an easy-to-use tool that will help them to explore unexpected differences between estimates and the actual outcome of the 2020 Census at the county level.[1]

2 Data

We obtained data from several sources to create a comprehensive dataset for this tool. The full list of files used and specific data cleaning steps can be found in Appendix A. We used two main files to calculate the difference between population counts and population estimates. The source data for the 2010 and 2020 census counts came from the Census Bureau's PL 94-171 file, also called the redistricting file (Manson et al., 2021). We obtained the group quarters counts from the 2010 Summary File 1 data, since the 2010 PL 94-171 data did not include group quarters counts (Manson et al., 2011). For the population estimates, We obtained the 2010 postcensal population and housing estimates from the U.S. Census Bureau's Evaluation Estimates website (US Census Bureau, 2021b).

One major data processing step was needed to link the census counts to population estimates. In order to make comparisons between the data, it was necessary to reconcile the race categories in the decennial census with race categories that appear in the data used to produce population estimates. Since the 2000 Decennial Census, respondents were able to self-report a single or multiple of the five Office of Management and Budget (OMB) race groups (White, Black, American Indian or Alaska Native, Asian, or Native Hawaiian or Pacific Islander). A final selection, "Some Other Race," also existed, and could be selected separately or in tandem with the other five race categories. Because the administrative data sources underlying the population estimates do not include this "Some Other Race" category, it was necessary to recode "Some Other Race" respondents from the decennial census data into the other five OMB categories (or one of the multiple race categories made up from those) so that the data frames were comparable. Typically, the Census Bureau calculated these apportionment ratios in a Modified Race file, but at the time of writing this paper, the U.S. Census Bureau has not created a Modified Race file for the 2020 Census. According to Jensen and Johnson (2021), the Census Bureau is likely to release a Modified Race file once the Demographic and Housing Characteristics file is published sometime in the late summer or early fall of 2023. Because of this gap in data availability, and the need to be able to compare race groups in the interim, we created our own modified race file. Unlike the demographers at the U.S. Census Bureau, we did not have access to individual-level

[1] We also developed metrics of census data quality at the block level through our Census Geographies Project.

data, so we used county-level aggregates with the following rules to create this file, based on the US Census Bureau's previous modification methodology from 2010:

1. For all people who did not respond with "Some Other Race," no changes were made.
2. For those who only reported "Some Other Race," we allocated these back to one of the five OMB race groups (or a combination thereof) using the national level proportion of people who were coded from "Some Other Race" to each of six race groups used in the population estimates for 2010: White, Black, American Indian or Alaska Native, Asian, Native Hawaiian or Pacific Islander, and Multiple Races. We then use controlled rounding, a raking procedure, to ensure our county total populations do not change when we undertake this recoding.
3. For people who responded "Some Other Race" in combination with one or more of the other five OMB race categories, "Some Other Race" was simply removed, leaving only the OMB race category responses (or the combination thereof).

After we reconciled the race categories, we joined the decennial census data with the estimates and calculated the raw and proportional differences between the estimates and decennial data.

3 Methods

The 2020 CCAT required two major methodological considerations: How to disaggregate the data so that local and state data users can understand how the counts compare across housing and demographic indicators, and how to define and measure what "divergence" is for the purposes of the tool.

3.1 Category selection

First, we chose which categories or variables from the census and population estimates to compare and how to compare them. We included total population and housing units to track broad trends in unchanging variables. According to Schneider (2021a), we knew that group quarters were particularly challenged due to the Covid-19 pandemic, so we included this variable. We also included persons per household and vacancy rates, which could provide additional insight into demographic changes. Jones et al. (2021) outlined how significant national discussion has centered around the shifting racial composition of the United States. To enable more detailed exploration of these patterns, we included the following race and ethnicity categories:

- White non-Hispanic
- Black non-Hispanic
- Asian non-Hispanic
- American Indian/Alaskan Native non-Hispanic
- Native Hawaiian/Pacific Islander non-Hispanic
- Two plus races non-Hispanic
- Hispanic (any race)

In all, twelve categories were included for analysis across population and housing characteristics. These categories were listed in the tool along with their divergence score, described in detail below.

3.2 Measuring divergence

To understand potential areas of concern for the 2020 Census, we calculated the divergence between the 2020 Census numbers and the Vintage 2020 population estimates. Traditionally,

Table 1: 2010 cut points examples.

Variable	Size	Count	Mean	1σ low	1σ high	2σ low	2σ high
GQ total	Large	578	−0.0274	−0.226	0.171	−0.425	0.370
GQ total	Medium	595	−0.000602	−0.515	0.514	−1.03	1.03
GQ total	Small	1970	0.444	−4.96	5.85	−10.4	11.3

this comparison—what demographers call the "error of closure"—is used to understand how to improve the population estimates. However, we used the comparison to understand potential decennial census issues given the high quality of the 2010 Census, the relatively strong approach to population estimation that the U.S. Census Bureau has created for the period between major decennial censuses, and the historic issues that the 2020 Census faced. The error of closure can always be caused by any one or a combination of three errors:

- Errors in the postcensal estimates base (i.e., the prior census, in this case the 2010 Census).
- Errors in the current decennial census (the 2020 Census).
- Errors in the components of population change that are used to move the population forward from 2010 to 2020.

Any analysis that looks at divergence between these two measures of the population should be used as a starting point, not a finish line. To better understand where the divergence between the 2020 Census and the Vintage 2020 population estimates was unexpectedly large, we used the difference between the 2010 Census and Vintage 2010 population estimates as the guidepost for defining what "expected" divergence looked like. We computed the actual 2020 divergence and the "expected" 2010 divergence by first using the PL 94-171 and estimates data from 2010 and 2020. We then calculated the proportional differences between counts and estimates in 2020 for each category measured and repeated this process using 2010 counts and estimates to establish cut points for "expected" divergence.

We then classified each county based on its population size. We labeled counties below 40,000 population as "small", counties between 40,000 and 100,000 population as "medium", and counties above 100,000 in population as "large." We calculated the mean and standard deviations (σ) by the county size and category to create the 2010 cut points (see Table 1). Then we compared the 2020 proportions to these cut points to establish whether the 2020 divergence was "close to expectation" ($<1\sigma$), "slightly divergent from expectation" ($>1\sigma$), or "highly divergent from expectation" ($>2\sigma$). This set of measures was selected for their simplicity and effectiveness.

Since the U.S. Census Bureau's estimates program does not create estimates of vacancy rates, we were not able to calculate the divergence for this variable. Instead, we calculated the vacancy percentage for 2010 and 2020 independently. When the difference between the 2010 and 2020 rates was greater than 5 percent, we considered the count to be "highly divergent". When the difference was between 2.5 and 5 percent, it was "slightly divergent". When the difference was less than 2.5 percent, it was "close to expectation".

4 Limitations

Any analysis of the potential quality or accuracy issues in a decennial census has limitations. Our comparison of the 2020 Census counts and the Vintage 2020 estimates can identify areas of unexpectedly high divergence, but it cannot identify the source of the error. Differences between the population estimates and the decennial census can be caused by many different types of errors, as laid out in the prior section.

Since our tool cannot identify the source of this error, it does not directly measure the quality of the 2020 Census. Our tool also does not determine if unexpectedly high divergence in a given area is due to error in the 2020 Census, in the Vintage 2020 population estimates based on errors in 2010, the components of change, or a combination of the three. Instead, we utilized indicators to identify areas of unexpected divergence and provide supplementary data to help users understand potential sources of error, particularly in the 2020 Census counts.

A large divergence alone does not necessarily mean the 2020 Census is either accurate or inaccurate, and agreement between the 2020 Census and the Vintage 2020 population estimates does not necessarily imply accuracy of the 2020 Census. We attempted to illustrate this visually by avoiding value-laden “good” and “bad” colors like green and red to illustrate divergence.

Our reliance on divergence in 2010 as our measure of “expected” divergence was a further limitation. Even though the 2010 Census was generally regarded as one of the most accurate in United States history, that does not mean that everything was accurate in that census. In fact, in subjects like group quarters or counts for some race groups, we find large divergence between the 2010 Census and the Vintage 2010 population estimates. For example, since group quarters in 2010 experienced a high level of divergence, similarly high levels of divergence in 2020 may be within expectation as we define it for the purposes of this tool, meaning a county classified as “close to expectation” may still have a high level of divergence.

Another limitation is centered around the data on race. Comparing race data in 2020 against 2010 is difficult for two main reasons. First, as we discussed earlier, the U.S. Census Bureau has not released a file on modified race that would allow us to more easily compare the 2020 Census to population estimates data. Because of this, we created our own modified race file. Although we know this solution has limitations (e.g. if the pattern of those who report “Some Other Race” has changed markedly since 2010), we believe data users need this data earlier than the U.S. Census Bureau can provide, and that our approach was a reasonable interim solution. Without reallocating the “Some Other Race” category, the decennial census data would not be comparable to the population estimates, and anyone who selected “Some Other Race” in the decennial census would not be represented in the tool.

Second, the U.S. Census Bureau changed their race coding in 2020 to be more inclusive of respondent intentions. This change led to more people being coded as multiple races than in the past and caused some comparisons for race to be harder to make between 2010 and 2020. While we acknowledge this is an issue in any race comparison for this decade, we think it is still an important task to undertake. We therefore caution users when looking at and interpreting the data, especially differences for the two or more races category, as can be seen in Figure 1.

5 Findings

The Census County Assessment Tool cannot definitively claim that any county was miscounted or experienced significant measurement error in the 2020 Census. However, it can enable the exploration of patterns across indicators and measurement categories. Using case studies in regionally and demographically different areas of the country, we can explore the tool results and generate research questions or hypotheses.

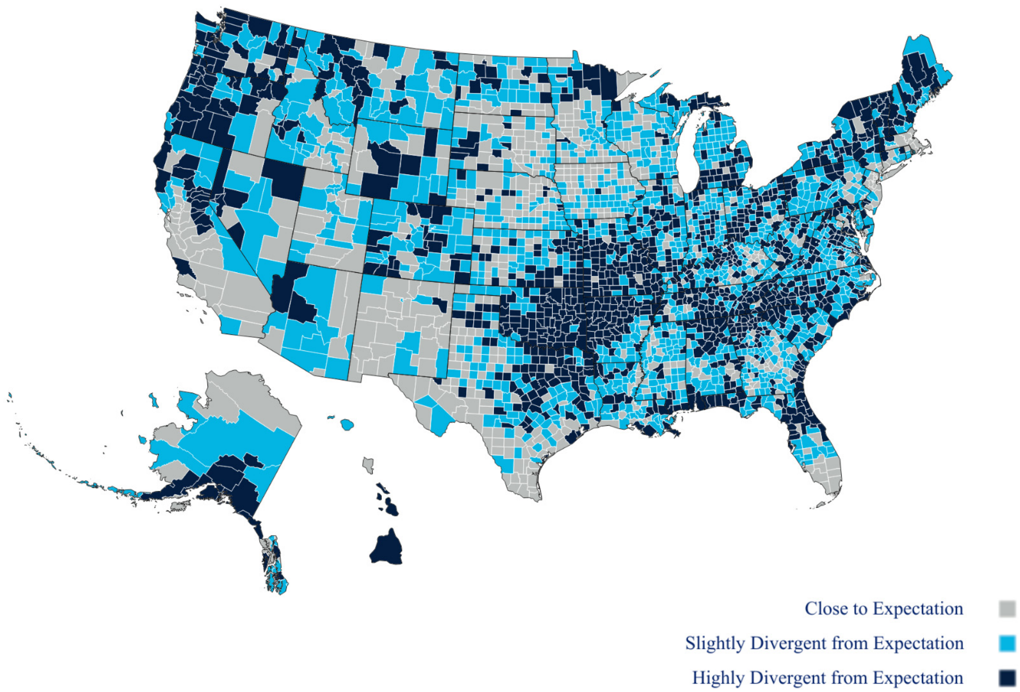

Figure 1: Two or more races by county.

5.1 Case Study: Madison County, ID

Madison County, Idaho is a small college town with a total population of 52,913 in 2020. It is the home of Brigham Young University-Idaho, a four-year university affiliated with The Church of Jesus Christ of Latter-day Saints (LDS). BYU-Idaho had 20,592 students enrolled and on campus in the fall of 2019 (Walker, 2019), and students are given latitude in deferring their education to go on missions (volunteer efforts to proselytize, offer church services, and provide humanitarian aid to specific countries and communities), so the overall student population may be even higher than 20k. According to internal church statistics, the LDS Church had 67,021 missionaries serving in an official, full-time capacity in December 2019, a portion of which were likely students at BYU-Idaho (The Church of Jesus Christ of Latter-Day Saints, 2020).

In Madison County, the total population count from the 2020 Census was 31.4 percent higher than the Vintage 2020 estimate, with 12,647 additional residents observed in the official redistricting count, making it "highly divergent from expectation" (see Figure 2). College towns experienced significant issues with 2020 enumeration (Schneider, 2021b). However, most college towns had precipitous drops in population due to college students dispersing during virtual semesters. Madison County is an anomaly for having such a dramatic *increase* in expected population.

One theory could be that the disruption of COVID-19 on missionary work may have brought a significant number of missionaries back to Madison County. Some 26,000 missionaries returned to their home countries in March 2020 (Stack, 2020). Certainly not all 12,647 additional residents in Madison County could be attributed to the return of LDS missionaries, but this could be a partial explanation of the divergence, and the group quarters count for Madison County was over 2,200 percent higher than the population estimates (see Figure 3).

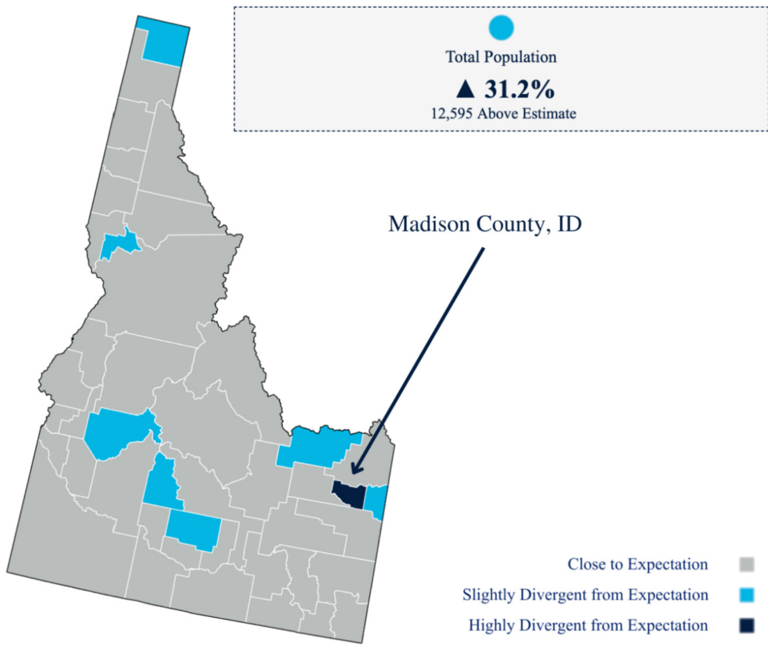

Figure 2: Madison county, Idaho.

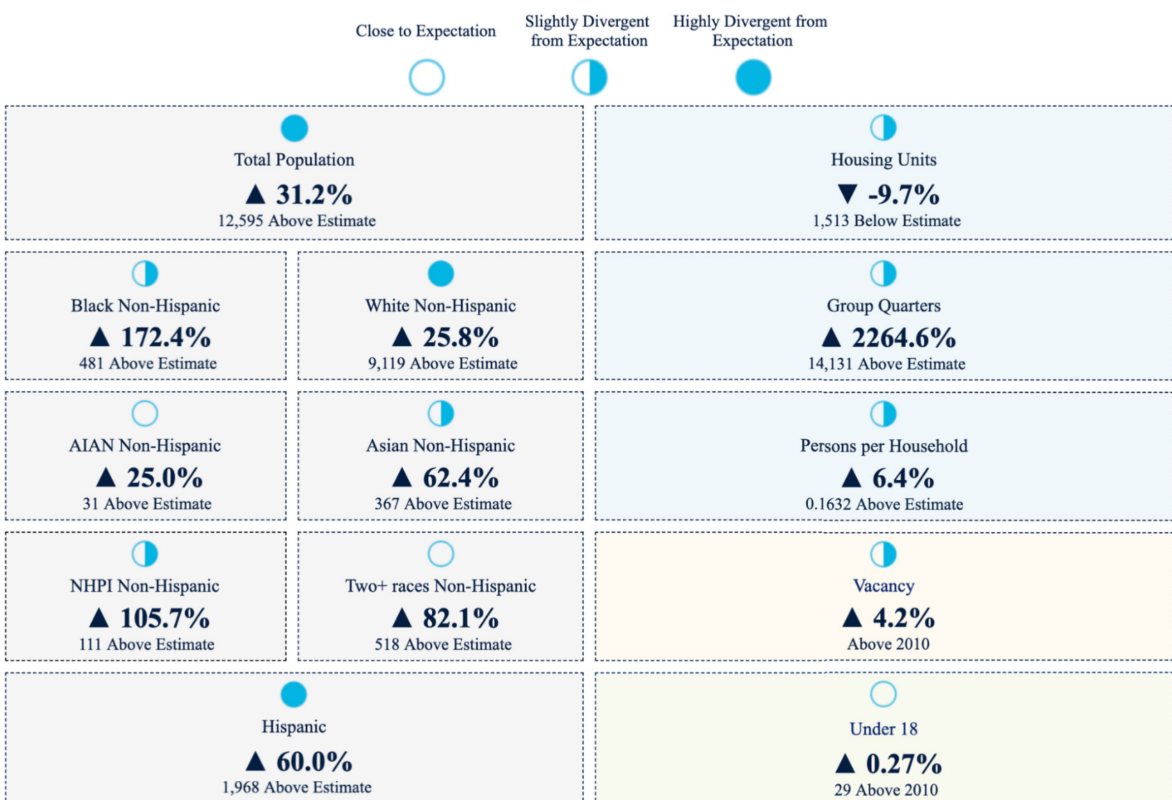

Figure 3: Measures of divergence in Madison county, Idaho.

Another source of potential error across demographic indicators, particularly for group quarters such as college dormitories, is imputation. The Census Bureau imputes missing data with data from similar nearby households. By the end of the census enumeration in 2020, the Census Bureau was missing as many as 1 in 5 group quarters counts (Schneider, 2021a). By looking at indicators of self-response rates, we can hypothesize what the impact of imputation was on the total population.

Madison County's self-response rate was higher than the national average by almost 5 percent. Additionally, when a university sends a verified list to the Census Bureau of all the students living on campus, this does not count towards the total self-response score, so the imputation rate could be even lower than what is suggested by the graph in Figure 4.

Since the 2020 Census was the inaugural internet-first census, internet access indicators are also important for understanding self-response. Madison County's broadband access was almost 7 percent lower than the national average, despite its self-response rates being higher. This condition could be a result of on-campus library computers or "Get Out the Count" efforts, but operational metrics such as the imputation rates for a county can help local governments and universities understand the accuracy of their census counts, especially for counties with high group quarters counts.

5.2 Case Study: Dallas County, TX

Dallas County, Texas was the home of over 2.6 million people in 2020. The Vintage 2020 estimates predicted more than 24,000 additional people in Dallas County in 2020, which is 0.9 percent of the total population. This difference is still "close to expectation". However, Harris County, Texas, the home of Houston, only experienced a count of 3,247 people below estimates despite being home to Houston, Texas, a city with a population twice as large as Dallas. Houston is also physically larger, with less population density.

Exploring the differences between these two counties may assist in identifying groups that

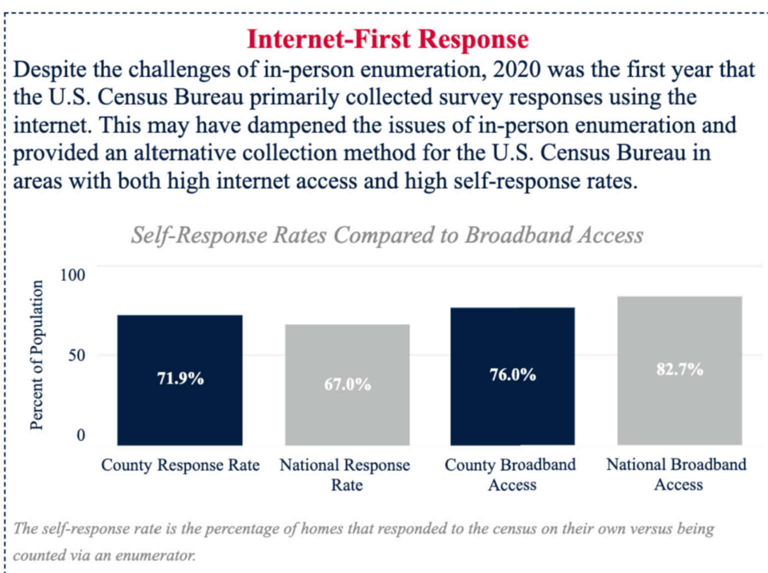

Figure 4: Response rates and internet access in Madison county, Idaho.

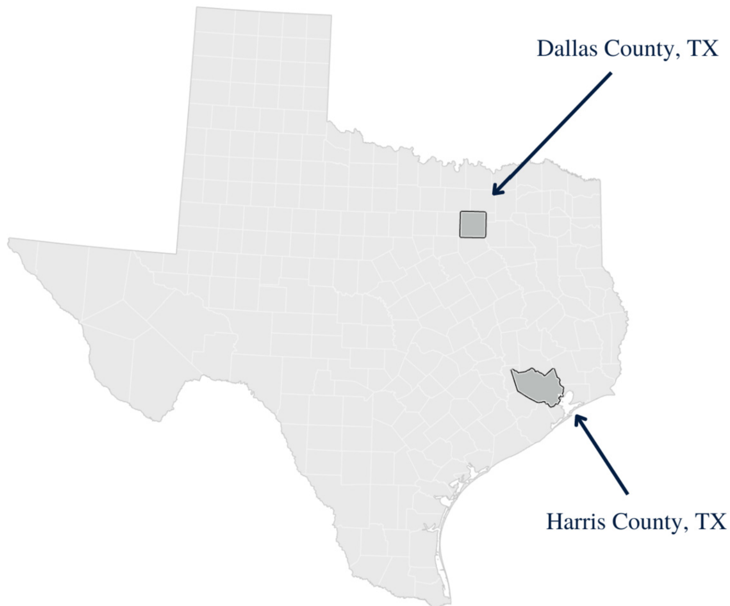

Figure 5: Harris and Dallas counties in Texas.

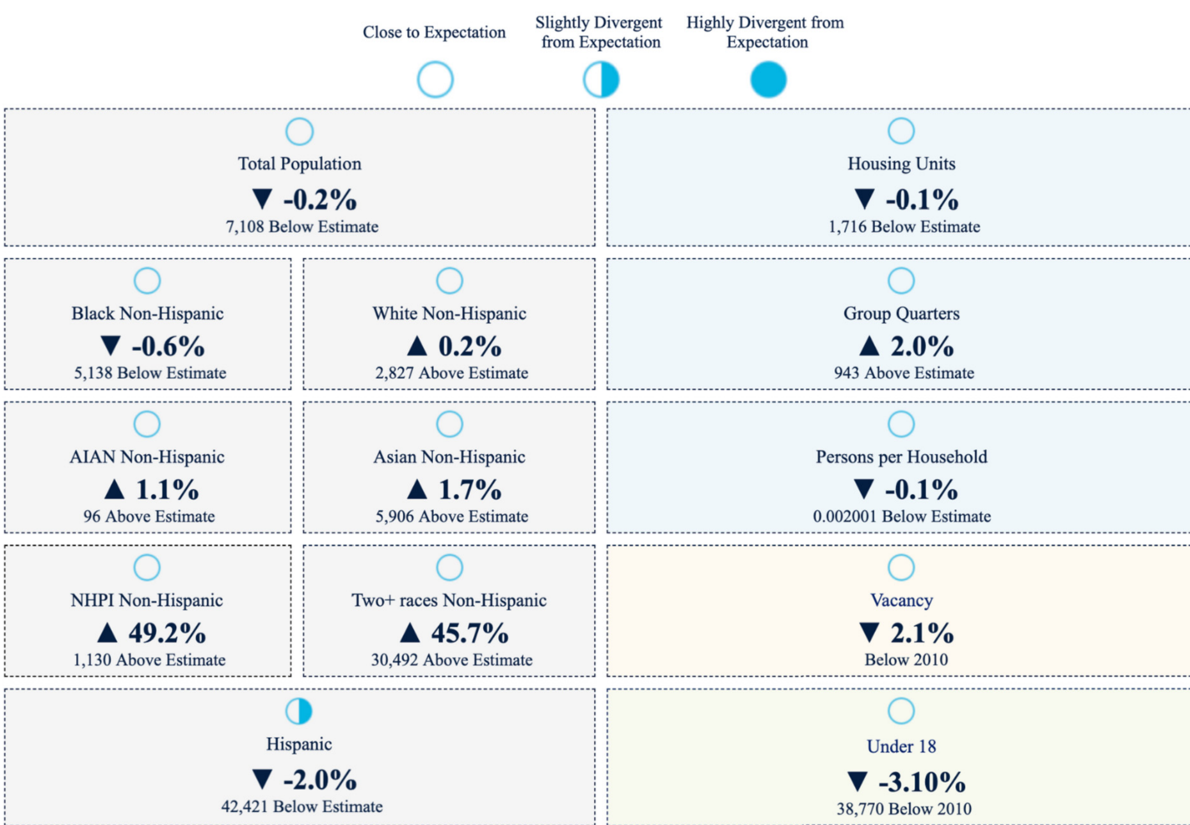

Figure 6: Harris county measure categories.

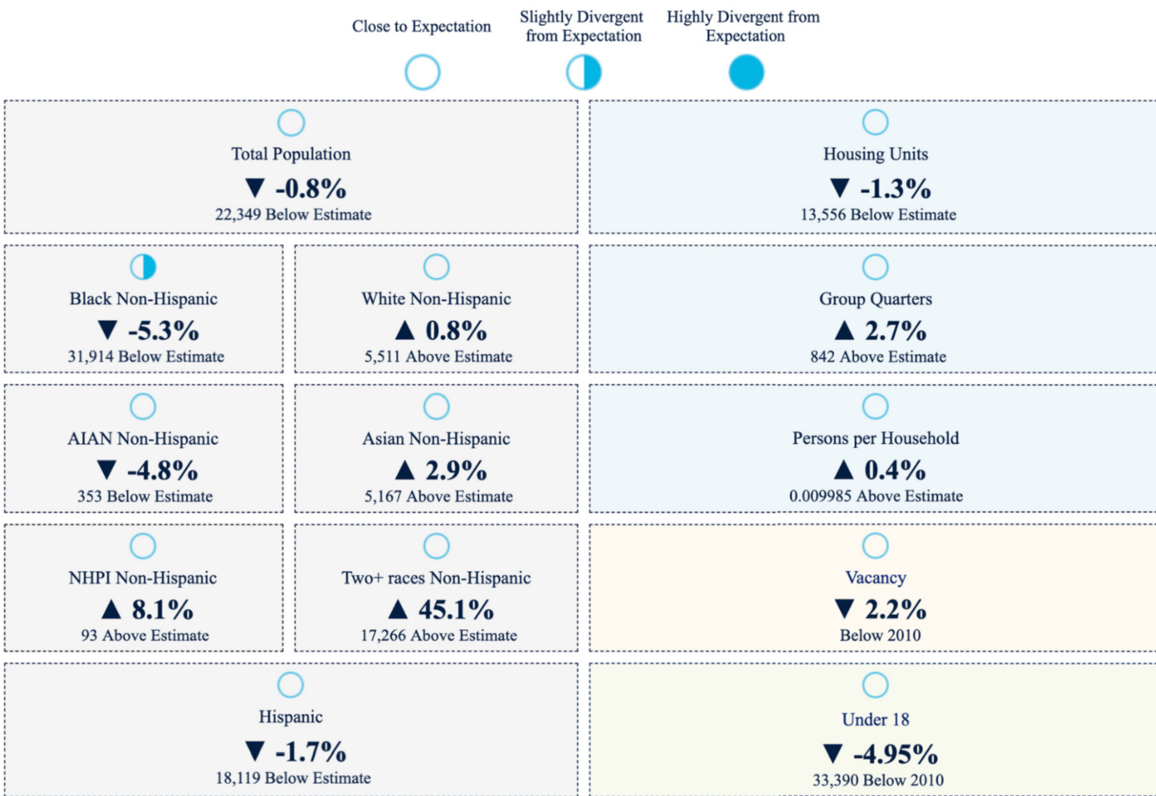

Figure 7: Dallas county measure categories.

could have experienced measurement error. Dallas and Harris counties had similar unemployment, COVID-19 case rates, self-response rates, and broadband access rates, and neither experienced a declared disaster. Therefore, we are limited to using the divergence measures by category to generate hypotheses to explain the differences in divergence rates.

Dallas County's housing units count was more than 13,500 below the Vintage 2020 population estimates, while Harris County's count was only 1,716 below the estimate. If Dallas County's rate of new housing construction was much slower than projected in the components of change, this could partially explain the difference in counts across person-level indicators.

Demographically, Harris County's Hispanic population was "slightly divergent from expectation", with the Vintage 2020 estimates predicting 36,600 more Hispanic residents than were actually counted. Meanwhile, Dallas County's Black, non-Hispanic population was also "slightly divergent from expectation" with 32,269 fewer residents counted than was predicted by the population estimates. Both counties experienced similarly divergent percentages of residents identifying as two or more races or Hispanic.

Overall, across most race and ethnicity groups, Dallas County performed worse than Harris County, and could point to where the hardest-to-count communities in Dallas may be, the difference in Get Out the Count funding (Harris County and the city of Houston invested over 5 million dollars together, while Dallas County and the city of Dallas invested about 1.9 million dollars total), or it may point to an error of closure within the housing component of change and downstream impacts on other measures. County officials in Dallas can test these theories to determine why these differences in divergence exist between Harris County and Dallas County.

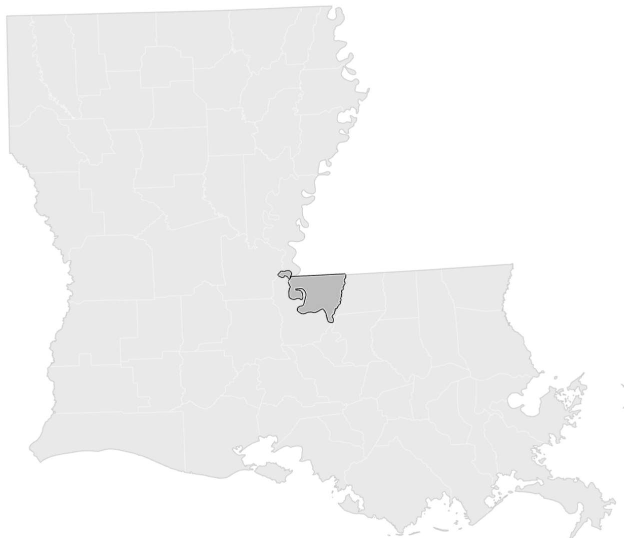

Figure 8: West Feliciana Parish, LA.

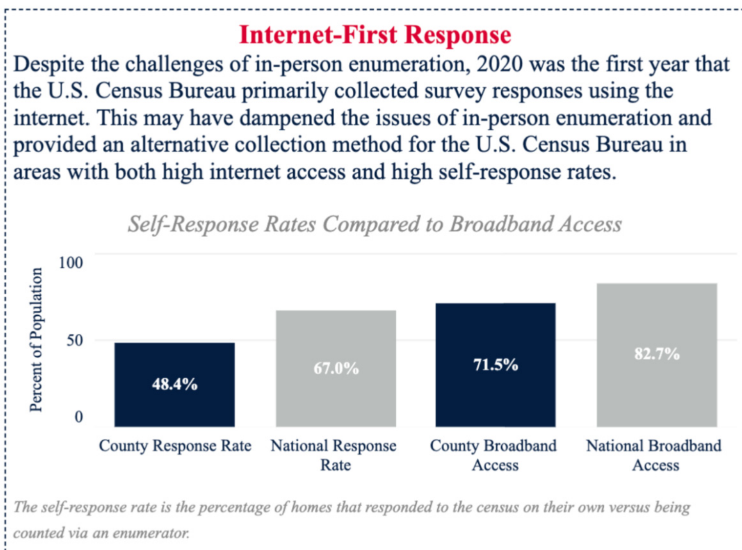

Figure 9: West Feliciana Parish self response rates.

5.3 Case Study: West Feliciana Parish, LA

West Feliciana is a largely rural parish in Louisiana with a total population of 15,310. West Feliciana is also home to the largest super-maximum security prison in the United States; the Louisiana State Penitentiary. This super-maximum prison has a total prison population of roughly 5,500 (Rubin et al., 2020). As such, at least one third of the parish resides in group quarters, which were particularly challenging to count in 2020. Similar to university dorms such as those in Madison, Idaho, the Census Bureau often relied on prison administrative reporting or imputation to count residents and determine their characteristics. This effect can be seen in the parish's self-response rates, which are significantly lower than the national average.

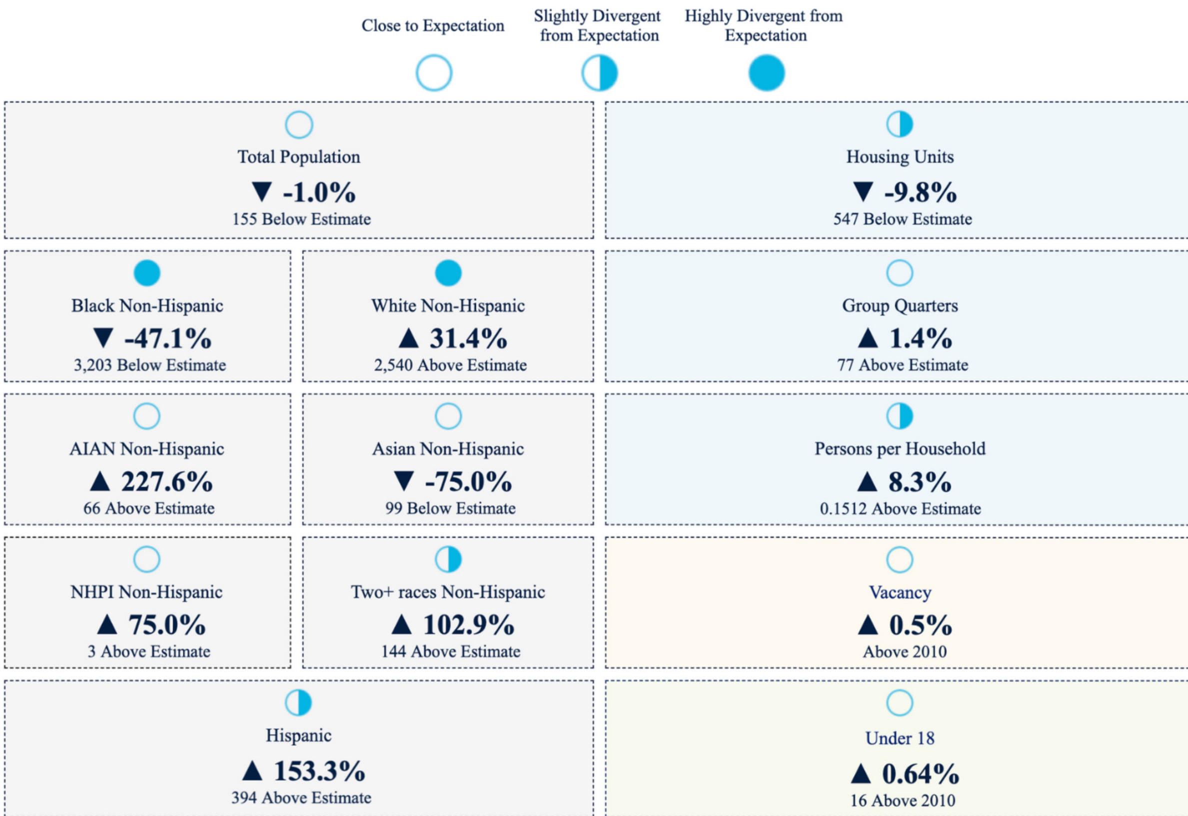

Figure 10: West Feliciana Parish measure divergence.

Between 75 to 97 percent of the population at the Louisiana State Penitentiary is unlikely to ever be released (Grob Plante, 2016). Despite the relatively static nature of the prison population, the measure categories for West Feliciana Parish show that the Black, non-Hispanic population was more than 3,000 residents or 47.2 percent lower than expectation, and the white non-Hispanic population was more than 2,500 residents or 31 percent above expectation. Furthermore, the Hispanic population was 159.4 percent higher than the Vintage 2020 population estimates. This kind of demographic shift is highly unlikely. If prison administrators either did not keep accurate records of prisoners' demographics, or if the Census Bureau imputed characteristics, this could have led to these dramatic divergences. Without more information on imputation rates for group quarters from the Census Bureau, we are unable to determine the underlying cause of the divergence.

In order to test the hypothesis that imputation was responsible for this dramatic change in population demographics, parish administrators can reach out to Louisiana's State Data Center or a Federal-State Cooperative for Population Estimates contact, both of which are listed in the CCAT. If these analyses result in discovering issues with the census counts due to imputation, parish administration can work with state contacts and U.S. Census Bureau staff to improve population estimates between censuses for their parish.

6 Discussion

The Census County Assessment Tool allows a user to explore various indicators of measurement error for the 2020 Census. While this tool does not enable a user to make a causal claim about undercounts or overcounts, it does point to areas for further analysis and contribute to the development of hypotheses which can be evaluated using high-quality research methods and data from administrative sources. The tool provides the contact information for state data stewards who may be able to assist users in understanding data quality issues from the 2020 Census.

This tool contributes to the field of data quality indicators for the 2020 Census by using methods similar to those already approved for data assessment by the U.S. Census Bureau (namely, Demographic Analysis). This tool allows state and local data users to finally have a resource for sub-state geographic indicators of 2020 Census data quality, and provides the data, code, and tool in an open, downloadable format for replication and analysis. The 2020 CCAT can be adapted to include additional measures as data are released by the Census Bureau. This tool begins to give local data users the kinds of metrics they need to develop high quality analyses for their jurisdictions.

Appendix A

We obtained the Federal Emergency Management Agency (FEMA) data from Disaster Declarations Summaries (Federal Emergency Management Agency, 2022). These records were subset to include only incidents that occurred in 2020 for all states. Outlying areas were not included. Incidents that occurred in 2020 but were not declared disasters until a later date were included in this dataset. Emergency Declarations were not included. County codes for American Indian/Alaska Native (AIAN) areas are not in the FEMA file; they are treated as independent entities by FEMA. Each declaration in an AIAN area was hand coded to its respective county or counties. In cases where one of the tribal or trust lands was included almost wholly within one county and an extremely small part of that area was in another county that had not already declared a disaster, the disaster was not attributed to that part of the county.

We acquired the decennial self-response rates from the U.S. Census Bureau's Decennial Census Self-Response Rates APIs (US Census Bureau, 2021a). The U.S. Census Bureau already performed transformations to collect the 2010 response rates in 2020 geographies, so the only transformation needed was to adjust the geographic boundaries. We selected the 2010 Final Self Response Rate along with the Overall 2020 Cumulative Self-Response Rate. We gathered data on the percent of county population with access to broadband from the American Community Survey 2015-2019 5-year data (Manson et al., 2020). Data on average weekly COVID-19 case rates came from the New York Times rolling case averages per 100k people from 2020. We downloaded the national and county-level averages from the New York Times GitHub repository (The New York Times, 2022). We used each Sunday's 7 day rolling average during 2020 as reference points for the average cases per 100k people.

We obtained each county's monthly unemployment rates from the Bureau of Labor Statistics' Local Area Unemployment Statistics (US Bureau of Labor Statistics, 2020b). We created GEOIDs from the series ID and transformed the period variable into month categories in the correct date format. We obtained national unemployment rates from the Bureau of Labor Statistics' Current Population Survey. We manually created this data frame using the "Labor Force Statistics from the Current Population Survey (not seasonally adjusted data, in thousands)"

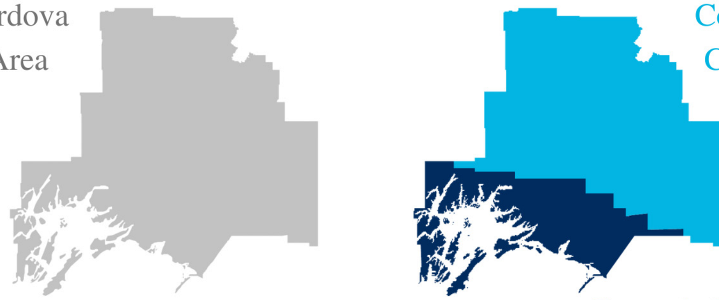

Figure 11: Alaska geographic changes.

table to find the national unemployment rate for each month in 2020 (US Bureau of Labor Statistics, 2020a).

To accommodate county boundaries and name changes between 2010 and 2020, specific recoding of files was needed. Bedford City, Virginia (51515) was absorbed into Bedford County, Virginia (51019). In 2015, Alaska changed the name and geographic identifier (GEOID) of the Wade Hampton Census Area (02270) to the Kusilvak Census Area (02158) and South Dakota changed the name of Shannon County (46113) to Oglala Lakota County (46102). For data products with a geographic vintage of 2010, Bedford City and Bedford County were combined and the GEOIDs of Kusilvak and Oglala Lakota counties were changed to conform to 2020 geographic identifiers.

In 2019, Alaska's Valdez-Cordova Census Area was split in two, creating the Copper River and Chugach Census Areas. In order to ensure the correct mapping of GEOIDs onto the 2020 data, an additional GEOID for the Valdez-Cordova Census Area was added to all data with a pre-2020 vintage. All data sources, regardless of the original geographic vintage, were transformed to include all three Alaska county-equivalent census areas (Valdez-Cordova, Chugach, and Copper River). A comprehensive list of the files used and their geographic transformations can be found in Table 2.

Funding

This project was developed in partnership with the Urban Institute with generous funding support from the Tableau Foundation and the Alfred P. Sloan Foundation. We are grateful to them and to all our funders, who make it possible for us to advance our missions. The views expressed are those of the authors. Funders do not determine research findings or the insights and recommendations of our team.

Table 2: Files used.

Data	**Source**	**Geographic vintage**	**Transformation**
Local Area Unemployment Statistics	Bureau of Labor Statistics	January 1, 2019	Chugach and Copper River Census Areas GEOIDs were added, county unemployment values were mapped directly from Valdez-Cordova
Public Law 94-171 Redistricting Files: 2010, 2020	U.S. Census Bureau	January 1, 2020	The sum or weighted average for Chugach and Copper River Census Areas were used to create the Valdez-Cordova Census Area
Evaluation Estimates: hu-est2010, hu-est2020, co-est2010-alldata, cc-est2010-alldata, cc-est2012-alldata, co-est2020-alldata, CC-EST2020-ALLDATA6	U.S. Census Bureau	January 1, 2010; January 1, 2020	The sum for Chugach and Copper River Census Areas were used to create the Valdez-Cordova Census Area
2015–2019 5-year American Community Survey	U.S. Census Bureau	January 1, 2019	Chugach and Copper River Census Areas GEOIDs were added, values from Valdez-Cordova Census Area were copied to these geographies
Decennial Census Self-Response Rates: 2010, 2020	U.S. Census Bureau	January 1, 2010; January 1, 2020	The Copper River Census Areas' response rate was used to create the Valdez-Cordova Census Area's response rate
Disaster Declarations Summaries	FEMA	January 1, 2020	There were no declared disasters in Chugach or Copper River; as such, Valdez-Cordova Census Area also had no disasters
COVID-19 Weekly Averages	New York Times	January 1, 2020	Copper River and Chugach Census Area both had no reported COVID-19 weekly averages; NAs were mapped to Valdez-Cordova Census Area

References

Rubin A, Golden T, Webster RA (2020). Inside the U.S.'s Largest Maximum-Security Prison, COVID-19 Raged. Outside, Officials Called Their Fight a Success.

Walker E (2019). Fall 2019 enrollment figures show continued growth.

Jensen E, Johnson S (2021). Using Demographic Benchmarks to Help Evaluate 2020 Census Results.

Federal Emergency Management Agency (2022). Disaster Declarations Summary.

Schneider M (2021a). Census: 1 in 5 dorms, prisons had no data at end of US count. *AP News*.

Schneider M (2021b). College towns plan to challenge results of 2020 census. *ABC News*.

Jones N, Marks R, Ramirez R, Rios-Vargas M (2021). 2020 Census Illuminates Racial and Ethnic Composition of the Country.

Stack PF (2020). Thousands of LDS missionaries are heading to their new assignments. *The Salt Lake Tribune*.

Grob Plante S (2016). Why the Inmates at Angola Prison Are So Invested in Their Craft Fair. *Racked*.

Manson S, Schroeder J, Van Riper D, Kugler T, Ruggles S (2011). IPUMS National Historical Geographic Information System: Version 16.0 Summary File 1. IPUMS, Minneapolis, MN.

Manson S, Schroeder J, Van Riper D, Kugler T, Ruggles S (2020). IPUMS National Historical Geographic Information System: Version 16.0 ACS 2015-2019. IPUMS, Minneapolis, MN.

Manson S, Schroeder J, Van Riper D, Kugler T, Ruggles S (2021). IPUMS National Historical Geographic Information System: Version 16.0 P.L. 94-171 Redistricting Data Summary Files. IPUMS, Minneapolis, MN.

The Church of Jesus Christ of Latter-Day Saints (2020). 2019 Statistical Report for April 2020 Conference.

The New York Times (2022). Coronavirus (Covid-19) Data in the United States. Original-date: 2020-03-24T23:41:39Z.

US Bureau of Labor Statistics (2020a). Labor Force Statistics from the Current Population Survey.

US Bureau of Labor Statistics (2020b). Local Area Unemployment Statistics.

US Census Bureau (2021a). Decennial Census Self-Response Rates (2020, 2010). Section: Government.

US Census Bureau (2021b). Population and Housing Unit Estimates Tables. Section: Government.

US Census Bureau (2022). U.S. Census Bureau Releases 2020 Undercount and Overcount Rates by State and the District of Columbia. Section: Government.