Title

Census

ISBN

9781555707378

Authors

Forte, Eric

Smith, Kelly

Sklar, Annelise

Publication Date

2011

Peer reviewed

Chapter 16

Census

Introduction

The decennial census and the United States Census Bureau were introduced in Chapter 10, but these resources are of such importance that they merit their own chapter. In many libraries, statistics are the most heavily requested genre of government information, and the decennial census is far and away the largest single statistical undertaking in the nation.

Think about the scale of the decennial census. Or better yet, think of the *second* largest statistical effort in the United States, the related American Community Survey (ACS)—which will also be covered in this chapter. The ACS surveys a whopping 250,000 households each month, and so reaches some three million households annually. This is much larger than any other government statistical survey, or any other regular statistical survey, for that matter. Yet the decennial census—by counting each and every person in the entire United States—counted some 310 million people in the year 2010, making it more than *100 times* bigger than the ACS.

The decennial census's core constitutional function is to determine the apportionment of representatives in the House of Representatives. But the census (especially now in combination with the American Community Survey) also provides the basic demographic and social statistics about our nation. It is the only nationwide statistical effort of sufficient scale to result in usable statistical data about even small cities and towns. In fact, this ability to access uniform, comparative data for these small geographies cannot be stressed enough. The size of the census, the data for small geographies, and the fact that it has been taken consistently for over 200 years

all make the census a singular and vital source of statistics.

Scientists, economists, and public policy makers use decennial census data heavily in research about the nation's populace. But perhaps the most important purpose beyond apportionment, and the one stressed in marketing campaigns designed to maximize voluntary census compliance, is census data's role in determining shares of billions of dollars in federal funds to states, cities, and communities. In fact, the Government Accountability Office (GAO) found that funding for all ten of the largest federal assistance programs--Medicaid, Highway Planning & Construction, State Fiscal Stabilization Fund-Education State Grants, Title I Grants to Local Education Agencies, Individuals with Disabilities Education Act Part B, Temporary Aid for Needy Families, Section 8 Housing Choice Vouchers, Community Development Block Grant, Federal Transit Formula Grants Programs, and the Children's Health Insurance Program--is determined largely based on census data (Orlando and Hyde 2010). For instance, when Medicaid funds are distributed to states by the federal government, they are distributed based on population figures and characteristics taken from the census. Census figures also helped determine shares of redevelopment money after Hurricane Katrina, and the same goes for dozens of other government programs. In short, the communities that most completely answer the census stand to win the most federal monies. In other words: when it comes to filling out the census, compliance literally pays.

The Census in the Constitution

While the Latin word *censere* more closely means estimate rather than count, the idea of a census has generally been associated with counting and especially counting people. Babylonians, Egyptians, Chinese, and indigenous Peruvians all appeared to conduct some type of census in

ancient times (Halacy 1980). In the Roman Empire, a census was often associated with either determining the numbers of men to potentially conscript for the purposes of war, or to count people for the purposes of taxing them. Neither reason is likely to have made a census a popular undertaking amongst citizens. Further, "the penalty for refusing to reveal how many people were in your household, how many slaves, how much livestock, was forfeiting it all and becoming a slave yourself" (Gibbs 2010:56). A census also appears several times in the Bible. Moses was told to conduct censuses of the Israelites, while King David is "said to have been punished for ordering a census—though the punishment was inflicted on the very people being counted (2 Samuel 24:15)," because it was Satan who tempted David into taking the census (Browning 2009). One gospel says that Joseph and Mary's compliance with the taxing purposes of a census was the reason they went to Bethlehem (Luke 2:1–6). While citizen concerns about participating in a census remain and will be discussed later, such concerns are now more likely to be for reasons related to privacy and accuracy rather than fear of conscription, taxes, or a vengeful god.

The United States' decennial census has a completely different mission than these ancient efforts, and its fascinating history goes far in explaining how we arrived at our modern decennial census. As introduced in Chapter 10, the decennial census is the original statistical program of the nation, drawing its statutory authority ultimately from Article I, Section 2 of the Constitution:

> Representatives and direct taxes shall be apportioned among the several states which may be included within this union, according to their respective numbers, which shall be determined by adding to the whole number of free persons, including those bound to service for a term of years, and excluding Indians not taxed, three fifths of all other Persons. The actual Enumeration shall be made within three years after the first meeting of the Congress of the United States, and within every subsequent term of ten years, in

such manner as they shall by law direct. (U.S. Constitution, art. I, sec. 2)

Debate in drafting this particular clause included various suggestions related to whether race or wealth was a factor in whether or not a citizen was even counted (Halacy 1980). The framers eventually intended slaves to be counted as only three-fifths of a person, which in addition to its shameful sentiment cost certain states representation. This practice lasted through 1860, when the Fourteenth Amendment took away the right to count any *male* person for apportionment purposes unless they had the right to vote, essentially ensuring that all inhabitants would be counted. (Although women were still counted in the census, it wasn't until the Nineteenth Amendment in 1920 that they were guaranteed the right to vote.)

The three-fifths clause was not the only portion of the founders' census article that was soon discarded. "Indians not taxed" referred to American Indians living free, a category that was largely gone by the end of the nineteenth century and officially retired in the 1940s. Indentured servants "bound to service for a term of years" similarly ceased to exist. The "direct taxes" line, meanwhile, was originally inserted to reduce temptation for a state to inflate its population in order to gain an unequal share of representation (for the greater the population, the greater the potential tax paid by the state). In practice, however, taxes were very rarely billed on the basis of census counts: only occasionally during early periods of wartime, although in the 1830s the federal government did distribute an *excess* of federal funds to states on the basis on census counts. The Sixteenth Amendment, in 1909, removed the concept of direct taxes entirely from the census, thus meaning that four facets of the founders' original census clause eventually disappeared from the law.

Census Privacy

All census data is confidential for 72 years, although it was not until 1840 that census takers were specifically instructed to treat all information with confidentiality (Halacy 1980). By 1880, confidential returns became law (13 U.S.C. 9). Census data may only be used for statistical purposes; no taxation, conscription, or other punitive measures are to result from filling out a census form. An individual or household's census answers are even protected from subpoenas or search warrants, so they cannot be used as legal evidence. This privacy requirement applies to workers in the field, those working with data at the Census Bureau, and the published public data; what an individual provides on a census form will be nothing more than anonymous numbers for most, if not all, of his or her lifetime. On a related note, citizens who do not answer the census can be fined, although in practice this has almost never happened.

Conducting the Census

The first census, taken in August of 1790, was conducted by U.S. marshals, who traveled the nation on horseback and simply went looking for people. With citizens not always expecting or wanting to be found, the enterprise carried some danger. While these marshals received instructions in how to conduct the count, there were no official forms, and marshals were responsible for procuring paper and organizing it in such a way as to facilitate a count. It was a labor-intensive process (see Figure 16.1).

<Insert Figure 16.1. Excerpt from 1790 Census Schedules Showing Name of Head of Household and Tick Marks for Each Resident of the Home>

Taking the census of an ever-expanding population remained a massive undertaking, especially as it continued to be taken by hand, in person. Although population growth made the

census increasingly difficult to conduct with each passing decade, improved methods somewhat mitigated the increase in volume. In 1840, the government finally instituted an official form for the marshals to use. The census of 1870 saw the introduction of the first rudimentary counting machines. In 1871, the census contracted with a young scientist to devise a better tabulation machine in time for the 1880 census. The scientist, Herman Hollerith, famously succeeded in building a machine that could rapidly read encoded punch cards (Figure 16.2). Hollerith and his invention later became the basis for the company IBM. The 1880 census also saw the U.S. marshals finally replaced by workers specifically dedicated and trained to conduct the census, creating the first large, once-a-decade class of temporary census workers. Counting was further eased in 1950 by the first use of an electronic computer to aid in tabulation. The 1960 census, meanwhile, saw the census recorded by the use of blackened dots which were microfilmed, and then read and tabulated by cameras and computers.

<Insert Figure 16.2. Punchcard from 1910 Census Used to Encode Each Household's Answers, Which Were Then Tabulated by Machine>

Modern-day census data is tabulated in the manner discussed in Chapter 10, with very large microdata files from which are made various specific reports and tables. Parts of contemporary decennial census microdata are available to the public, allowing researchers to create their own tables and reports out of the original data. See the section on Decennial Census Microdata later in this chapter for more information.

Amazingly, all of these censuses were still conducted in person—a huge undertaking with the now extremely large population. The 1910 census saw the first effort at a mail-in census; response was abysmal, and the effort was abandoned. The 1960 census again saw forms mailed

to households. Households could not mail them back, but at least the forms would be filled out and ready to share when the census worker came knocking, saving time and expense. Finally, in 1970, the government stopped sending workers to every household; citizens were able to mail back census forms, and only nonrespondents received a personal visit.

It's natural to wonder about the viability of taking an online census. Like online voting, fears about the integrity and security of such an effort long kept an online census from occurring. In 2010, the most recent decennial census, forms were once again sent and returned via U.S. mail, with nonrespondents receiving a personal visit or phone call. In 2013, however, recipients of the American Community Survey were finally given the option of responding online. Census 2020 is expected to be conducted primarily online, largely to secure cost savings and increase response rates (http://www.census.gov/2020census).

Census Content

The content of the questions is the single most important aspect of the census. If a question was asked as part of a decennial census, a wealth of data exists on that topic for the entire nation. The questions form the basis of available data and publications.

Almost from the beginning, the census evolved from the simple headcount called for by the Constitution. As early as 1790, James Madison proposed using the opportunity of the census effort to gather further information. Despite some opposition—these early efforts were denounced by some as "a waste of trouble and supplying materials for idle people to make a book" (Cassedy 1969, 216)—Congress soon acted on the potential to learn more about the nation via the census effort. Questions related to age, sex, ancestry, schooling, and wages presented a broader demographic portrait of the nation, while questions about the economic structure helped

Congress understand the fast-evolving state of the nation's industry and agriculture. Over the ensuing years, these extra questions have come and gone. In addition to many economic and demographic questions, the census has made some unique and interesting inquiries. For many years, people were asked about their place of religious worship, which aroused suspicion and some controversy in a nation built on religious freedom and individual privacy. (Law now forbids any government survey of religious worship, practice, or beliefs.) Citizens were asked how many and what type of bathroom facilities they had (i.e., indoor vs. outdoor). People were asked to record how many insane or "idiotic" persons lived in the household (presumably, the insane were not the ones filling out the forms) as well as habitual drunkards and tramps (Halacy 1980). Regardless of whether the questions make sense today, what had begun as a simple census had grown over a century into a census of hundreds of questions about health, education, housing, and ancestry, with even more separate questionnaires for American Indians, the recently deceased (filled out by surviving family members), and soldiers, among others.

The invaluable census publication *Measuring America: The Decennial Census from 1790 to 2000* includes a history of the census along with copies of every question asked and every form used in each census. Especially useful is a grid (excerpted in Figure 16.3) that summarizes the subjects asked over each census, making it easy to see, census by census, if a particular piece of data was gathered.

While the specific questions asked shifted over the years, extra questions remained a staple of the census in one manner or another up through the 2000 census. The 2010 census, for reasons discussed shortly, saw these extra questions largely jettisoned and a return to what is the shortest and simplest census since the very first one in 1790.

Statistical Sampling

Asking the extra questions, as important as they were, eventually did begin taking a toll, as the more expansive census became both harder for citizens to answer and more expensive for the nation to conduct. Picture a census worker sitting in a household and asking over 100 questions. And then going next door. And so on, through the entire nation.

The mid-twentieth century saw two changes to address these issues. First, much of the extra data that the census had been gathering began to be covered by other surveys. For instance, a great deal of health and education data are now gathered by much smaller surveys (see Chapter 10). Economic questions, which comprised many of the extra questions, were spun off into a separate economic census (covered in Chapter 15). Finally, the extra questions that remained—and there were still dozens not related to the basic headcount used for apportionment—began to be asked of only a sample of the population, taken concurrently and as part of the census. In this scenario, 100 percent of households were still counted and asked basic questions such as age and sex; however, a smaller sample of the population would answer a census that included the extra questions. Since this extra data was not used for the constitutional apportionment, a statistical sample was and remains sufficient to gather this data, as modern statistical techniques ensure an accurate count of these questions.

<**Figure 16.3.** Excerpt of Grid Summarizing All Questions Asked in the Decennial Census from 1790 to 2000>

Beginning in 1940 and continuing through Census 2000 (changes for the 2010 census are discussed below), approximately one in five to one in six of all households received the longer sample form, which included the basic questions and also the extra questions. This became

known as the census long form. Everyone else received the census short form. The data from those who filled out the short form (plus the short form elements of those who filled out the long form) is referred to as 100 percent data, to make clear that this data is from all respondents to the census, unadjusted by sampling. The actual percentage of people receiving the long form varied by geography. Part of what makes the census so invaluable is its ability to provide data for even small geographies such as small towns. In a big city, one in six is still a massive sample and very statistically accurate. In a small city, though, asking only one in six does not yield enough data. A larger portion of the population in smaller places and rural areas were given the long form in order to make the results statistically significant; as many as one in every two households in the smaller geographies were given the longer survey.

The content of these extra questions is always subject to debate and sometimes controversy. The Census Bureau receives thousands of suggestions for questions to include on the census, most of them related to a specific interest of a researcher or organization. For instance, if one is selling shoes, what better way to collect data about shoe usage or preferences than through the census? It's paid for by taxes and gathers data for even small geographies, so a shoe company could easily see which cities and towns bought which types of shoes. There is nearly no end to what people hope could be asked in the census, from the serious to the silly: "As long as we're spending all this money to reach so many people, imagine what we could find out. Which do you favor, Leno or Letterman? Smooth or chunky? Faith or works? Liberty or equality?" (Gibbs 2010, 56). The Census Bureau, however, rejects nearly all such requests as outside the scope of the decennial census.

Census Controversies

The census has been central to numerous formative political and social controversies in American history. Simply figuring out how to use the population data to determine the size of the House of Representatives was contentious right from the first census effort. How big should the House be, and most importantly, what is the population threshold for a new representative? Disagreement over how to determine the allotment of representatives led to the very first presidential veto in the nation's history when George Washington vetoed the Apportionment Act on April 5, 1792.

The years leading up to the Civil War saw the census at the center of a number of very divisive arguments related to the balance of power between slaveholding and free states. In the published reports of 1860, the Census Bureau editorialized about what the numbers meant with regard to slaves and their potential freedom, suggesting that some combination of morality, barbarianism, and genetics would see the black population soon become either extinct altogether or totally absorbed into the white population—an example of official government racism (Kennedy 1864). Less controversial but more famous, in 1890 the Census Bureau declared the closing of the frontier as shown in the latest census figures.

Contemporary census controversies are both familiar and novel: familiar in that the census still scares some part of the population, who fear it as an intrusion into their privacy and a manifestation of a big brother government; novel in that the census's consistent inability to accurately count every person has raised the idea of using more sophisticated statistical means to make a more accurate census.

Privacy advocates have long resisted the extra questions on the census, arguing for keeping the census to only its strict constitutional role of counting inhabitants and apportioning

representatives. Even the 2010 census—one of the shortest in history—had its detractors. Minnesota Representative Michele Bachmann, perhaps the most vocal of 2010 census critics, claimed at the time to be "leaving most of its form blank except for the question that asks directly how many people reside in her home. 'We won't be answering any information beyond that,' she said. She complained that the remaining nine questions are 'very intricate, very personal' and argued that 'the Constitution doesn't require any information beyond that'" (Colvin 2010:41). Echoing concerns about government intrusion into private information, Representative Jeff Duncan (R-S.C.) and several co-sponsors introduced the failed "Census Reform Act of 2013," which would have abolished virtually all census surveys except for a simple decennial count of the population. Despite such resistance, voluntary census response rates have remained steady over recent decades, with over 70 percent of households returning the form via mail and the rest receiving personal visits.

While the law suggests that these privacy concerns are unfounded, census data has been abused in the past. The most notable instance was during World War II, when the Census Bureau provided the War Department names and addresses of Japanese Americans and thus aided efforts to identify Japanese Americans for internment (Holmes 2000). Scholars Margo Anderson and William Seltzer, who discovered and reported this abuse of census data, have done extensive work related to government statistics and confidentiality (see https://pantherfile.uwm.edu/margo/www/govstat/integrity.htm).

The other great contemporary census controversy surrounds the problem of undercounting the population. Poor areas and racial minorities are especially prone to being undercounted, a phenomenon first truly noted when more African Americans registered for the World War II draft than were even thought to be in the country based on census data. As Holden

summarizes, statisticians have long advocated for using statistical sampling to adjust the results to more accurate figures (2009). While sampling has been used for the extra questions, it had never been the basis for the apportionment data, which is taken from 100 percent results only. To address this in time for Census 2000, the Census Bureau proposed a plan to use limited statistical sampling to make population numbers more accurate, arguing that such methods were sound and fair and would represent a more accurate count. Before such a plan could be instituted, however, the Supreme Court, in the 5–4 decision *Department of Commerce v. U.S. House of Representatives* (525 U.S. 326), ruled sampling for apportionment purposes to be unconstitutional.

So to make the census as accurate as possible, the Census Bureau relies on vigilance, education, and outreach rather than statistical methods. It tries hard to make people understand that by not responding to the census, they only hurt themselves and their communities and they are essentially going unrepresented in the House of Representatives. To maximize an accurate count, the census even has a large-scale marketing campaign, which in 2010 saw the first ever census Super Bowl commercial. Successful outreach can be a difficult proposition, though. The homeless population, for a number of reasons well-articulated by Kearns (2012), is historically difficult to reach and enumerate. Undocumented immigrants, legally included in census enumerations, often have a distrust of the government that leads many to be fearful of providing personal information via census forms. Despite the efforts to ensure participation, the 2010 Census still missed millions of people, especially African Americans, Hispanics, and Native Americans.

Race, ethnicity, and sexual orientation are other often contentious issues in the decennial census, and changes in policy have made these demographics difficult to accurately compare

over time. For example, the first time that respondents were able to identify themselves as more than one race was in the 2000 census. O'Hare (1998) provides background and summary of federal statistical classification of race categories, including the change for Census 2000. This move has been both welcomed and criticized, and is indicative of ongoing sensitivity in the nation to issues of race. Ancestry questions (which ask respondents to identify their country of origin), meanwhile, are consistently dogged by claims of undercounts and concerns that respondents do not always willingly identify their ancestry. Complicating matters even more, studies have shown that many people are inconsistent in the way they identify their race and/or ethnicity (Cohn 2014). It wasn't until the 2010 census that the Census Bureau, at the direction of the Obama administration, counted and reported same-sex married couples. In 1990, when the "unmarried partner" option was first included to count unmarried heterosexual couples living together, the Census Bureau - because same-sex marriage was not yet legal in any states - edited answers of self-identified same-sex spouses so that one member of the couple appeared to be of the opposite sex. In 2000, the Census Bureau did the opposite, re-categorizing couples who identified as "married" as "unmarried partners" (Long 2011).

Census Geography

Geography is key to the decennial census. Obviously, the census must count the people in each state in order to apportion representatives by state. But in order to distribute federal funds—not to mention for the benefit of research about the nation—the census also presents data that can describe smaller geographies, such as cities and towns. The census does that and more, providing data for states, counties, cities and townships, and smaller geographies known as census tracts, block numbering areas, block groups, and blocks. This last geography is more or less akin to an

actual city block. All of these smaller geographies, from census tracts on down, began in the twentieth century, and it was not until the 2000 census that every place in the country was assigned a census tract. Previous decennial censuses, then, do not feature census tract (or equivalent) data for all places.

Figure **[originally known as 8.2 (see Chapter 8, p. 197)]** showed major census geographic entities; Figure 16.4 shows the census geographies that a particular address belongs to.

Census Geography, Major Divisions, Values for Address Shown in Figure 16.4

1. Census Region (of four regions: Northeast, Midwest, South, West): West
2. State: California
3. County or equivalent: Santa Barbara County
4. County subdivisions: n/a
5. Place (city or equivalent): Santa Barbara city
6. Census Tract (within a county, between 1,200–8,000 people; ideally 4,000): 5.02
7. Block group (within census tract, a cluster of blocks of 600–3,000 people): 4
8. Census Block (within block group, much like an actual city block): 4,000

<Insert Figure 16.4. On the Top, Census Map of Santa Barbara, CA; on the Bottom, Inset Showing Census Tract 5.02; Block Group 4; and Block 4000>

While beyond the scope of this chapter, the Census Bureau also developed the TIGER (Topologically Integrated Geographic Encoding and Referencing) system, geographic reference files that, in addition to their use in census geography, helped spur the development of the world of GIS data and software. These TIGER files were created by combining United States

Geological Survey topographical maps with Census Bureau maps of addresses used by decennial census takers in the field. When used with GIS software, TIGER files combined with decennial census data enable creation of sophisticated maps, including features such as roads, rivers, elevation data, and legal and census boundaries, all mapped with demographic and socioeconomic data from the census (see Figure 16.5 for an example).

Census Publications

Although neither the constitution nor later action by Congress called for publication of census results, publications (thankfully for librarians and for this chapter) did start with the first census, with a single 56-page report. In 1800, another single report was produced, of 74 pages. By 1840, there were well over 1,000 pages of decennial census reports; by 1880, 20,000; by 1960, over 100,000 pages (Eckler 1972). That's hundreds of times bigger than *War and Peace*. And the 1970, 1980, and 1990 censuses were much bigger than that, with hundreds of published volumes. As of 2000, most data and publications began to be available via the web as part of the same rapid transition to web access that has transformed so much government information. Within one ten-year decennial census period, access evolved from these hundreds of printed reports (with limited and difficult-to-use electronic access via CD-ROMs and primitive online systems) to a full point-and-click web interface.

<Insert Figure 16.5. Map Showing Average Commute-to-Work Times, by County, in the Northeast Corridor: The Darkest Areas Commute over 30 Minutes; the Lightest, under 11>

Early Census Reports: 1790–1840

Using the earliest decennial census reports is straightforward. As published in print or scanned online by the Census Bureau (http://www.census.gov/prod/www/decennial.html), these historical

reports usually contain a summary table or two and several pages of data per state. The state data is further broken down by smaller geographies, namely, by county and city/town/place, and occasionally, a subdivision of a place, such as Harlem Division in Figure 16.6.

The published data depend on the questions asked. For each relatively simple decennial census in this early period, there is only a volume or two of results: one volume of the basic population data by geography, and sometimes a second volume or two for extra data such as the economic or pensioner data. From 1790 to 1840, the census generally just took a count and gathered data by sex and a few basic age categories, and for free citizens versus slaves. The 1810 census saw a concurrent effort to count manufacturers, the result being some data on the number, volume, and type of manufacturing done in the nation, states, and for each county. Separate volumes were published to share this data. The 1820 census was the first to add an economic question as part of the census itself, with citizens noting whether they were employed in agriculture, commerce, or manufacturing. Again, a second volume was published to present this data. The 1830 and 1840 censuses more closely resembled the early censuses, although 1840 included some extra questions asked of military pensioners.

Accessing the data in these volumes is pretty straightforward. There is usually a summary page with general population data for each state, and then a chapter for each state with further breakdowns by county and town. This is important as it mirrors—in a simpler form—all census data throughout history, namely, with data gathered by geography.

Not all reports are online at the historical census reports website. The invaluable *Catalog of United States Census Publications 1790–1972* lists and describes every report in each census. Getting one's hands on these volumes is not always easy. While historical census reports have sometimes been republished by private publishers, it is mostly larger libraries that own

significant runs of historical decennial census reports. The Census Bureau also has a print-on-demand program for reports that it has scanned. Finally, each state is home to a State Data Center (http://www.census.gov/sdc) that coordinates access to census data. Each of these sources can aid users in actually laying hands on the appropriate decennial census volume.

<Insert Figure 16.6. 1790 Census Summary Results for the Nation (Left) and Part of New York (Right)>

Multivolume Reports: 1850–1930

Beginning in 1850 the amount of data gathered in the decennial census grew quickly and steadily. This was partly because more extra questions were asked and partly because of the nation's rapidly expanding population. The year 1850 was notable in that it was the first census to ask additional questions not only related to manufacturing and industry but also about education, literacy, and real estate, among others (Figure 16.7). The 1880 census was notable as the first to ask questions related to health and health conditions. And 1910 saw the first use of census tracts and smaller geographic designations, for the first time allowing users to access data about particular parts of towns. Use of census tracts expanded over the following decades, mainly in metropolitan areas, but eventually the Census Bureau expanded the program to include the entire county by 2000.

<Insert Figure 16.7. Page from 1850 Census Volume on North Carolina Showing Literacy Data by City and Occupations by State>

By 1880, the decennial census had become undeniably long and the number of published reports numbered into the dozens. Extended social and demographic questions were asked, as well as many questions about agriculture, industry, and manufacturing. By this time finding the

desired data becomes more complicated. Different volumes cover different geographies and different topics. Data for a particular place might appear in many different volumes, depending on the subject; similarly, data for a particular subject might appear in many different volumes, depending on geography.

To find specific data, consult the *Catalog of United States Census Publications, 1790–1972*, mentioned in the previous section. Like census reports from earlier eras, some of the reports from this period will be a part of the Census Bureau's historical census reports website; others will be owned by depository libraries, and perhaps State Data Centers.

Era of the Statistical Sample: 1940–2000

As noted, the 1940 census was notable for the introduction of statistical sampling. Census workers—still on foot, remember—asked an extended set of questions to every twentieth household. Sampling continued to be refined over the coming decades. With the growth of the nation and the census—data was also being tabulated for thousands of smaller geographic units, such as census tracts—published reports by this time number in the hundreds for each census, and finding the desired data can take some digging. In 1970, published reports also began to be broken out by whether the data was from the 100 percent questions or the sample questions, which is occasionally helpful in identifying the right volumes. For instance, there are single volumes of the basic 100 percent data for each state (covering places and smaller geographies within the state, too); there are also single volumes for each state with the sample data results, and volumes containing data by census tract for metropolitan areas. Many subject-based reports have been produced, too. In addition to the listing of all reports in the *Catalog of United States Census Publications, 1790–1972*, each recent census has a separate guidebook. And to get a

sense of the scale of decennial census publishing, there is also the (now discontinued) annual *Census Catalog and Guide*, which lists the available reports from each recent census (see Figure 16.8). For 1990, just the *listing* of decennial census publications spanned over 30 pages.

Once in a volume, the Table Finding Guides (Figure 16.9), which are just after the table of contents, are quite useful. With their grids of subjects by geographies, these guides allow users to quickly find the right table.

Schulze's three-volume set (1983, 1985, 1988) is also very useful in navigating the library of decennial census volumes from the first decennial census through 1980, using handy grids to identify years, volumes, subjects, and geographies, and lead to detailed listings of data by volume.

The most recent census data are best—and increasingly only—accessed online via the American FactFinder, which provides simpler discovery of data by subject and geography. However, because of "the extremely large file sizes, American FactFinder can only store data from the last two Censuses," (U.S. Census Bureau Frequently Asked Questions) meaning that the 1990 decennial census data was dropped from American FactFinder when 2010 data was released. Print volumes may be found using the same method as for various older volumes: online as part of the historical census reports website; in local depository libraries; and at State Data Centers.

End of the Long Form and Birth of the American Community Survey: Census 2010

The 2010 census is the shortest and simplest census since the very first one. There is no long form and nearly no extra questions: nothing about education, earnings, occupation, transportation, physical characteristics of the dwelling, ancestry, languages, or other detailed

socioeconomic and demographic statistics. There is no sample of the population answering additional questions. Everyone gets only the short form, which counts the population, asks about relationships in each household, and gathers data about race and age. The results will form the basis of reapportioning Congress; be used to distribute shares of billions of federal dollars; and allow researchers, policy makers, and the public to access basic statistics about the populations of states, cities, counties, and smaller census geographies such as census tracts, block groups, and blocks.

<Insert Figure 16.8. First Page of 30 Listing Reports and Data Products of Census '90, from the *Census Catalog and Guide*, 1995>

The shift to this throwback census is due to the emergence of the American Community Survey (ACS). Because recent decennial census efforts saw increased difficulty in getting households to fill out the long and cumbersome long form, the Census Bureau—after years of planning and testing—instituted the ACS to replace the long form. The goal of the ACS isn't only to make the decennial census easier to gather, but also to provide ongoing current data.

<Insert Figure 16.9. Table Finding Guide from a Volume of the 1990 Census>

The ACS gathers demographic, social, and economic data similar to that previously included on the census long form. Age, race, sex, family relationships, housing characteristics, education and literacy, health and health insurance, employment and income, language, and transportation are all among the subjects of questions asked. Since the ACS is conducted continually instead of once a decade, it provides more current data than the census long form did.

In the ACS, approximately 250,000 households are sampled each month across the nation. This sample is large enough to provide estimates for larger geographies—states, cities,

and counties with a population of 65,000 or more— every year instead of every ten years, as the number of responses to these approximately three million surveys per year will yield statistically valid results. Smaller geographies require more time to accumulate sufficiently sound ACS data: three years for places with 20,000 to 65,000 people; and five years for places with fewer than 20,000 residents. As such, the ACS produces three distinct data sets each year: one-year data sets, three-year data sets, and five-year data sets.

This process—taking multiple years to gather enough data for smaller places—makes the ACS different than the long form in one significant way. Except for those places with over 65,000 people, which yield yearly data, ACS data is based on moving averages (also known as period averages). This means that if one is looking for data about Gunnison, Colorado, whose population is under 20,000, ACS will have data for the 2005–2009 period, because it takes five years of surveys to gather enough responses for an adequate survey of Gunnison. It will also have data for 2006–2010; 2007–2011; and so on. It will not have data for any single year because not enough residents of Gunnison will receive an ACS in a year to yield useful data. It is, therefore, important to understand that the resulting data is for a period of time, unlike data from the decennial census, which is for a single point in time.

These moving averages do occasionally present confusing options when using the ACS. All places with over 20,000 people will have multiple ACS figures available. For instance, Syracuse, New York (with a population well over 100,000), will have ACS data for the year 2007, the years 2007–2009, and the years 2005–2009. Each figure will be slightly different. All three are legitimate and useful, one just needs to understand the peculiarities of these moving averages. This is especially tricky when measuring something like prices, which can fluctuate over even short periods. (The Census Bureau is using inflation data to attempt to normalize such

statistics.) Herman (2008) offers an excellent overview of the issues surrounding the ACS in general and moving averages in particular.

Online Census Data and American FactFinder: 2000–2010

As seen, the Census Bureau has scanned many of its historical census reports, although nowhere near comprehensively, and made them available online at http://www.census.gov/prod/www/decennial.html. Also worth noting is the Historical Census Browser (http://mapserver.lib.virginia.edu/), a grant-funded project of the University of Virginia Geospatial and Statistical Data Center that uses an online interface to present a huge amount of historical census data, from 1790 through 1960, although only for counties and larger geographies. The National Historical Geographic Information System (http://www.nhgis.org/) provides a similarly interesting compilation of historical and current census data with advanced geographic capabilities.

Census 2000 was the first census that used the web as the primary and original data distribution method, via a portal called American FactFinder (see Figure 16.10). The American FactFinder (http://factfinder.census.gov), which was re-designed to coincide with the release of Census 2010 data, is the primary freely-available method of accessing contemporary census data. It presents data about the population and economy of the United States and places within: it's simple to enter a state, city, county, place, or even a street address to get data for large or small geographic areas. Likewise, it's fairly straight-forward to find and generate tables by navigating through topics using either the advanced or guided search.

<Insert Figure 16.10. American FactFinder Guided Search Screenshot - http://factfinder.census.gov/faces/nav/jsf/pages/guided_search.xhtml>

The three largest data sources found in American FactFinder are the decennial census, the ACS, and the economic censuses, but it has grown to include the American Housing Survey, additional annual economic surveys from the Census Bureau, the Census of Governments and Annual Survey of Governments, and vintage Population Estimates. If one seeks population or demographic data for a particular state, county, or place, data is available from both decennial censuses and the ACS as well as the other included surveys. For instance, in Figure 16.11, data may be accessed for a particular place for either decennial census data or from the Population Estimates program. Options on the left allow for easy access of data by other demographic and socioeconomic characteristics, surveys, topics, and more.

While American FactFinder remains the main source for recent census data, there is a separate page for Census 2010 (http://www.census.gov/2010census/). Census 2000 and Census 1990 both also have their own gateway pages, pointing to the data and relevant documentation:

http://www.census.gov/main/www/cen2000.html

http://www.census.gov/main/www/cen1990.html

<Insert Figure 16.11. Using American FactFinder Advanced Search: Note the Options at Left - http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t>

Decennial Census Microdata and Enhanced Census Data-Based Products

Because advanced researchers often need microdata they can tabulate to their own specifications, the Census Bureau makes available Public-Use Microdata Samples (PUMS) for both the decennial census and ACS. These datasets are anonymized samples with records for housing units with individual responses characteristics. PUMS for 1980-2010 decennial censuses and the ACS are available through the Census Bureau's FTP site (http://www2.census.gov/). 2000 and

2010 decennial census and ACS PUMS are also available in American FactFinder by searching the "topic or table name" for the term PUMS. The ACS PUMS are also available via the Census Bureau's DataFerrett (http://dataferrett.census.gov/) data analysis tool. The Integrated Public Use Microdata Series (IPUMS-USA) (https://usa.ipums.org/usa/) is a free tool that makes available PUMS data back to the 1850 Census. Historic PUMS microdata is also available for download from the Inter-university Consortium for Political and Social Research (ICPSR) (http://www.icpsr.umich.edu/).

Decennial census data is also among the most fertile basis for enhanced, privately produced statistical and mapping tools. Since government statistical data is—within the limits of privacy law—public, third-party developers are able to use the data as the basis of new products. While such reuse is at times somewhat disingenuous—some publishers simply take government documents and reprint them under their own imprint, with perhaps only a new preface as added value—statistical data, especially from the decennial census, has seen some wonderful private development. Census data is behind countless products designed for market research (see Chapter 15, Business, Economic, and Consumer Information). For more in-depth research use, companies like Social Explorer (http://www.socialexplorer.com), SimplyMap (http://www.geographicresearch.com/simplymap), and PolicyMap (http://www.policymap.com/) use the data to create specialized thematic maps beyond what is possible using American FactFinder's mapping options. In 2012, the Census Bureau began providing access to datasets by API, and the website now includes an app gallery to highlight tools, created by the Census Bureau and by third parties, that utilize it.

Conclusion

The unmatched size and scope of the decennial census, and the related American Community Survey, make it the most important statistical repository from the government and the most widely-used single statistical undertaking in the nation. The decennial census continues to evolve, and understanding and finding data sources for any particular decennial census requires some research. Even when the reports and data are online, it is necessary to know the nature of the statistical effort in order to interpret the results. From American FactFinder to census microdata to printed reports, the Census Bureau houses an astonishing amount of statistical data about the nation, available to use in ways designed for the novice user through the statistician. The ACS, meanwhile, is a groundbreaking effort to provide updated demographic information even for small geographies.

Exercises

1. Explore the 1830 census. What statistics are available on race and slavery? Are there statistics by state? By city?

2. Find the population of Phoenix, Arizona, in every census from 1920 to 1990. What further geographic breakdowns exist by decade? When were census tracts introduced?

3. Generate a list of the commuting times to work for all cities in California from Census 2000.

4. For your current address, identify the census tract, block group, and block. What streets or geographic features form the boundaries for your block? What is the population, by age and sex, for your block?

5. What is the most recent median household income for the county you currently reside in?

The one you grew up in? For each, what is the specific source of the data?

Sources Mentioned in This Chapter

Sources mentioned in this section do not duplicate the References that follow.

1990 Census, http://www.census.gov/main/www/cen1990.html.

2000 Census, http://www.census.gov/main/www/cen2000.html.

2010 Census, http://www.census.gov/2010census/.

2020 Census, http://www.census.gov/2020census.

American Community Survey, http://www.census.gov/acs/www/.

American FactFinder, http://factfinder.census.gov/.

Bureau of the Census. 1974. *Catalog of United States Census Publications, 1790–1972*. Washington, DC: Government Printing Office. http://catalog.hathitrust.org/Record/000732965.

Bureau of the Census. *Census Catalog and Guide*. Annual, 1946–1998. https://www.census.gov/prod/www/catalogs.html.

DataFerrett, http://dataferrett.census.gov/.

Historical Census Browser, http://mapserver.lib.virginia.edu/.

Historical Census Reports, http://www.census.gov/prod/www/decennial.html.

Integrated Public Use Microdata Series (IPUMS-USA), https://usa.ipums.org/usa/.

Inter-university Consortium for Political and Social Research (ICPSR), http://www.icpsr.umich.edu/.

National Historical Geographic Information System, http://www.nhgis.org/.

PolicyMap, http://www.policymap.com/.

Public Use Microdata Samples, http://www2.census.gov/.

SimplyMap, http://www.geographicresearch.com/simplymap.

SocialExplorer, http://www.socialexplorer.com.

State Data Center Program, http://www.census.gov/sdc

References

Anderson, Margo J. 1988. *The American Census: A Social History*. New Haven: Yale University Press.

Browning, W. R. F. 2009. "Census." *A Dictionary of the Bible*. New York: Oxford University Press. Oxford Reference Online.

Cassedy, James. 1969. *Demography in Early America: Beginnings of the Statistical Mind*. Cambridge: Harvard University Press.

Cohn, D'Vera. 2014. "Millions of Americans Changed Their Racial or Ethnic Identity From One Census to the Next." *Pew Research Center FactTank*, May 5. http://www.pewresearch.org/fact-tank/2014/05/05/millions-of-americans-changed-their-racial-or-ethnic-identity-from-one-census-to-the-next/

Colvin, Jill. 2010. "Down for the Count." *First Things* 202: 39–42.

Eckler, A. Ross. 1972. *The Bureau of the Census*. New York: Praeger.

Gibbs, Nancy. 2010. "Count Me In." *Time*, April 5, 56.

Halacy, Dan. 1980. *Census: 190 Years of Counting America*. New York: Elsevier/Nelson.

Herman, Edward. 2008. "The American Community Survey: An Introduction to the Basics." *Government Information Quarterly* 25 (3): 504–519.

Holden, Constance. 2009. "The 2010 Census: America's Uncounted Millions." *Science* 324 (5930): 1008–1009.

Holmes, Steven. 2000. “Report Says Census Bureau Helped Relocate Japanese.” *New York Times*, March 17. http://www.nytimes.com/2000/03/17/us/report-says-census-bureau-helped-relocate-japanese.html.

Kearns, Brandon. 2012. “Down for the Count: Overcoming the Census Bureau's Neglect of the Homeless.” *Stanford Journal of Civil Rights & Civil Liberties* 8: 155-182.

Kennedy, J. C. G. 1864. *Population of the United States in 1860: Compiled from the Original Returns of the Eighth Census under the Direction of the Secretary of the Interior*. Washington: GPO.

Long, Quaid. 2011. “Queering the Census: Privacy, Accountability, and Public Policy Implications of Adding Sexual Orientation and Gender Identity Questions to the US Census.” *DTTP, Documents to the People* 39 (4):15-18.

O’Hare, William. 1998. “Managing Multiple-Race Data.” *American Demographics* 20(4): 42–45.

Orlando, Lucia, and Rebecca Hyde. 2010. "Census 2010: What's The Big Deal?" *Dttp: A Quarterly Journal Of Government Information Practice & Perspective* 38(1): 10-12.

Schulze, Suzanne. 1983. *Population Information in Nineteenth Century Census Volumes*. Phoenix: Oryx Press.

Schulze, Suzanne. 1985. *Population Information in Twentieth Century Census Volumes, 1900–1940*. Phoenix: Oryx Press.

Schulze, Suzanne. 1988. *Population Information in Twentieth Century Census Volumes, 1950–1980*. Phoenix: Oryx Press.

U.S. Census Bureau, Frequently Asked Questions-., https://ask.census.gov/faq.php?id=5000&faqId=9222.

U.S. Census Bureau. 2002. *Measuring America: The Decennial Census from 1790 to 2000*. Washington, DC: Government Printing Office.