Finding the right demographic data

**Places to get data:**

**American FactFinder:** [**http://factfinder.census.gov/**](http://factfinder.census.gov/) This is the data-cutting tool on the Census Bureau’s web site. When you first pull it up you have 4 options – Community Facts, Guided Search, Advanced Search and Download Center. It’s ordered “easiest” to “hardest”. FactFinder includes all of the American Community Survey results from 2005 to present, plus the decennial censuses of 2000 and 2010. This site (especially the Advanced Search part) is very rich in terms of how much information is here, however that also makes it intimidating and not intuitive for the beginner. The Community Facts section, however, is a very quick and easy way to get a single statistic about a place.

**CensusReporter.org:** <http://censusreporter.org/> This site was built by journalists for journalists (they got a Knight Challenge grant to build it) and it continues to be upgraded. Its biggest weakness is that you can only get the most current American Community Survey data; you can’t look back in time, at all. However, it has some really big upsides that you can find on the Census website. First, the site is easier to navigate. Second, you can more easily compare different places. Third, you can view the data you’ve chosen either as a map, a table with whole numbers or a table with percentages. Finally, you can download the data in a variety of ways, including a shapefile that can be used in GIS software.

**NHGIS.org:** <http://www.nhgis.org> This is a site run by the Minnesota Population Center. It’s considered the go-to source for historical census data. You can find data from the very first census ever taken in the U.S., and everything up to present day (including the stuff that is on the other two sites I mentioned). Biggest downside: the user interface is clunky and it’s not super easy to pull out the data you want. Other thing you note: you need to have a username and login (it’s free) and you have to promise to cite your data source properly and to “use the data for good, never for evil.”

**See “other resources” section** below for other local sources that might have some stuff already pulled out.

**Which data source should I use?**

If you want….

* Current snapshot of one place on any topic (i.e. you want the total population, or the racial breakdown, or the percent over age 65 or the percent who are disabled)
  + Quickest option: American FactFinder, either Community Facts or Guided Search
  + Other option: CensusReporter.org
* Current snapshot of more than one place, particularly if you want to download the results into a spreadsheet and/or want to make a map
  + Best option: CensusReporter.org
  + Second option: American FactFinder Guided Search
* How something has changed over time in one or more places….
  + If you are okay only going back as far as 2000: American FactFinder Advanced Search
  + If you need to go further back in time: NHGIS.org

**Key concepts:**

**Which dataset** – Every 10 years, the Census Bureau conducts a **decennial census** in an attempt to count everybody. You can get data from each of these going back decades and decades, however the questions that get asked each decade tend to change. For example, the 2010 census only asked a few key questions – far, far fewer questions than what were asked in 1990 and 2000. For about the past decade, the Census has also been conducting annual surveys to get more frequent data on a wider variety of topics. This is called the American Community Survey. You will be far more likely to use ACS data if you want current data.

**Geography** – Everything is based on geography. So you have to first decide what level of geography you want data for. There are many options. Here are the big ones: nation, states, counties, congressional districts, census tracts, census block groups, census blocks. Census does say you can get zip code level data (they call it Zip Code Tabulation Area – ZCTA), but the boundaries of those zip code areas don’t always match the U.S. Postal Service’s actual zip codes.

Census does not generate data for what we consider “neighborhoods” in Minneapolis and St. Paul, however it is possible to combine census blocks, block groups or tracts to generate neighborhood-level data ourselves. [Minnesota Compass](http://www.mncompass.org/profiles/neighborhoods/minneapolis-saint-paul) has already done that using 2010 data.

Also note: The U.S. govt considers the Minneapolis-St. Paul metropolitan area to be much bigger than most of us realize, and any Census data from 2014 for the MSP metro would encompass: Hennepin, Ramsey, Dakota, Washington, Carver, Scott, Sibley, Mille Lacs, LeSueur, Chisago, Isanti, Sherburne, Wright, St. Croix WI, and Pierce WI.

**Universe** – The data you get will either be based on “people” or “housing units.” For example, we can find out how many people in Minnesota are of Hispanic ethnicity and that table would be based on people. Then there’s also a table about owner-occupied versus renter-occupied where the householder is Hispanic. However, the counts are based on housing units. So it’s showing the number of housing units – headed by a Hispanic person – that are owner-occupied. Regardless of which website you get data from, the heading or table title should tell you what the “universe” is. It should say people or housing units. Or it might say “people age 16 and older” or it might indicate that the table is only including people of a particular race.

**Topic crosstabs** – All of the data comes as a pre-formulated summary or a cross-tabulation. So your options are dictated by which cross tabulations the Census has chosen to provide. For example, from table B05003B (sex by age by nativity by citizenship status for Black or African American Alone) we know that about 28% of the black population in Minnesota was born in another country (foreign born). However, it’s almost impossible to find any other information about that foreign born subset of the black population (like what percent are in poverty, or what the median household income is for that group or what percentage are in the labor force, etc). You can find those things for the entire Minnesota foreign born population or the entire black population, but not for that subset.

**Lingo** – Probably the hardest thing to deal with is the lingo that comes along with Census data. The Census uses specific terms for things that you and I would label some other way. For example, if you want to find out how many single-parent households there are, or what the median income is for those households, you have to know to look in one of the cross tabulations related to “household type”. But there’s also “family type”, which is a subset of household type. Or let’s say you want to find out how many people were born in another country, then you have to look for “nativity.” But there’s also “ancestry” which is different (for example, I would say my ancestry is primarily German). I’ve provided you another handout -- Census-speak: A guide to the perplexed, by Paul Overberg – that does a great job of helping you navigate this.

**Margin of error** – With the American Community Survey data, it will include a margin of error number with your data. For example, it might tell you that there are 100 black people in Arden Hills and then it says the margin of error is +-20. That means that the number of black people is somewhere between 80 and 120 people. The margin of error is going to be higher when you “slice” down to smaller groups. So for example, let’s say we want to look at just the Asian people who live in Lino Lakes and make more than $50,000 in income. We have now sliced the population three ways and with each slice, you can expect a higher and higher margin of error. It also gets higher when the population you are measuring is small to begin with. You want to make sure the margin is not too big. For example, if the estimate is 100 and your margin is 50 or more – that would be too problematic. The other thing you need to be cautious of is if you are trying to compare two places or compare change over time – you need to make sure the difference is big enough to get outside the margin of error.

**ACS 1-year, 3-year and 5-year** – The American Community survey data is published in three ways, as 1-year data, 3-year data and 5-year data. The 1-year is data compiled from just a single year of surveys. The 3-year is from 3 years’ worth of surveys, etc. This makes it possible to provide more statistically sound data for smaller geographies where very few households would be surveyed in just that 1 year or 3-year span. Which one you use is primarily dictated by the level of geography you want. In fact, the 1-year data only includes results for geographies that have 65,000 or more people. For us that means you can get the state of Minnesota, the MSP metro as a whole, a handful of counties and just a few cities. However, the margin of error won’t be as strong as you’d get if you used the 3-year or 5-year. For very small geographies, like census tracts or blocks, your only option is the 5-year data.

Other Resources:

**Minnesota State Demographic Center** – <http://mn.gov/admin/demography/>, The state demographer, currently Susan Brower, and her deputy, Andi Egbert, are super helpful and do a great job of explaining this complicated stuff in an easy-to-understand fashion.

**Minnesota Compass**, <http://www.mncompass.org/> a publication by the Wilder Foundation, does a lot of work with Census Data and they would be a good place to check with to see if they have already published what you’re looking for; or they would also be a good place to seek help. Craig Helmstetter does their data work and always seems willing to help. As noted earlier, they publish great neighborhood profiles on Minneapolis and St Paul neighborhoods.

**Met Council**, <http://www.metrocouncil.org/Data-and-Maps.aspx>, They have a team of data people and they publish reams of data, including a lot from U.S. Census. Beware, though, that they publish their own population estimates (based on their own work) and you will find these differ from U.S. Census estimates. Both are worthy of using, you just need to be clear with readers where you got the data from.

**U.S. Census Bureau Public Information Officers** – another resource that can help you find the right table to use, <http://www.census.gov/newsroom/about.html>.

*@MaryJo Webster*[*mjwebster71@gmail.com*](mailto:mjwebster71@gmail.com) *May 2016*