

Census-speak

A guide to the perplexed

Paul Overberg

USA TODAY

poverberg@usatoday.com and @poverberg

Prepared for NICAR annual conference

Baltimore, Feb. 27-March 2, 2014

If you're just starting to use Census Bureau data, you may have struggled with the vocabulary. There's good [documentation](#) for the American Community Survey, but it's scattered and it typically just defines concepts, like "household" or "householder." You need to learn those before you can understand the terms of art – the data yardsticks that make Census data useful. This handout tries to offer a crash course in both. First, a narrative of very basic **concepts**, in **bold**. Categories in those concepts appear within "quotes." Then you'll find some yardstick terms at the end.

Concepts

Start with the most basic idea: a **housing unit** is a building where people can live. It can be **occupied** or **vacant**. If it's occupied, it's a **household**. If it's home to two or more people related by blood, marriage or adoption, it's the most common type of household: a **family**. (About 66% of households are families of some kind or other. The rest are **non-family households**. They include the 28% of households that are people living alone and the 6% are made up of "unrelated individuals.")

Whoever fills out the Census form in a household is by definition the **householder**. All relationships are defined with reference to this person. ("What if my visiting mother-in-law fills it out? Doesn't that mess things up" No big deal in the overall scheme of things.)

Most relationship categories are self-evident. A basic division: **relatives** and **non-relatives**. The former seems simple – people related to the householder by blood, marriage or adoption. But non-relatives span everything from "roomer/boarder" – often a financial arrangement – to "roommate/housemate" – an affinity relationship – to "unmarried partner" – a singular and often long-term relationship. Adding to the confusion: Same-sex couples can be legally married in some states, but even if they specify that one is a "spouse," their relationship will be recoded to "unmarried partner" on federal surveys. A federal interagency group is studying how best to update this critical question in the face of legal same-sex marriage.

Among relatives, one key concept frames many Census tables: **own children**. They are the never-married kids of the householder who are under 18. Unfortunately, counting adult kids of the householder isn't as easy and must be done indirectly. There's also no easy way to count some relationships that have bloomed in the increasingly complex U.S. household. Examples: "civil union," "sandwich parent," "boomerang kid," "live-in caregiver," "only child," "blended family."

The Census Bureau does tally **multigenerational households**, which it defines as three generations living in a household, but it doesn't typically break them down by their many types.

Some family terms aren't what they seem like. For instance, a "female householder – no husband present" living with her kids isn't necessarily a "single parent" (a term you won't find in Census tables). Why? Her boyfriend may be living there and doing a lot of parenting. As long as they aren't married, he's an

“unrelated individual” and hard to track in Census reports. He could even be the father of the kids. In another scenario, the woman’s mother or brother might live with her and her kids.

And what’s a **subfamily**? It’s a second family in the housing unit, like when your brother and his wife move into your basement after they lose their house to a tornado, economic or real.

People living in a housing unit have two kinds of **tenure**: **renters** or **owners** (either outright or under a mortgage). This has nothing to do with the building’s structure. For the last half-century, about two-thirds of U.S. housing units have been **owner-occupied**. **Physically**, two-thirds have also been **1-unit detached**. There’s a large but not complete overlap between those two groups. Both are ebbing slowly. On the rise: renters and **1-unit attached** housing units. (Think rowhouses or townhouses.) Don’t confuse these with **condominium status**, an ownership form that can apply to any type of housing unit.

Key to defining a housing unit: rooms intended for use just by their occupants with a door to the outside, or at least a corridor to one. This is to distinguish from **group quarters**, where 3% of Americans live. Most common: dorms and prisons. Also common: barracks, jails, halfway houses, nursing homes, monasteries, convents, etc. Assisted-living facilities sit in the middle of the spectrum. They are classified based on rules about how much their occupants share facilities.

When it comes to identifying people, **race** and **ethnicity** are two of the most basic ways. In federal statistics, they are two separate ideas, a split that causes no end of confusion. Race is the major way we define ourselves based on appearance. In some form or other, it’s been used since the first census in 1790. The Census Bureau uses the [rules](#) set for all federal agencies by the White House Office of Management and Budget. They have an imposing name: Statistical Directive 15.

The rules evolve. A 1970 addition: a separate concept called Hispanic ethnicity. You answer “yes” or “no.” It’s not just another answer to the race question. So you can be Hispanic and white, Hispanic and black, Hispanic and Asian, etc. Or non-Hispanic and white, non-Hispanic and black, etc. Many categories are sparse in many areas so they’re lumped together in many reports. As diversity grows, however, they become more meaningful in more areas.

This twofold system probably is due for overhaul. Tests during Census 2010 suggest that using a single question would produce better data more cheaply. It could also allow detailed race reporting that is now available to Asians (“Chinese,” “Filipino,” etc.) and American Indians (“Navajo,” “Iroquois,” etc.) and Hispanics (“Mexican,” “Salvadoran”, etc.) to be extended to whites and blacks. This would yield better data on new immigrants from the Middle East and Sub-Saharan Africa. Watch for debate and possible changes in the next few years.

There is a third way that the Census Bureau lets people express cultural identity: **ancestry**. It offers free-form entries and has programmed its computers to recode them into [hundreds](#) of national and ethnic groups. It publishes totals for more than 100 in standard tables. “American” is increasingly popular, but the largest groups remain European nationalities, such as “German” and “Irish.” Most are foreign, but a few are located in the United States, such as “Cajun” or “Tejano.” For Americans of European descent, this is often a “roots” question. The race question doesn’t offer detailed reporting for people who mark “white” or “black,” so the ancestry question is a way to find the origins of new immigrants from the Middle East (“Iraqi”, “Jordanian”, “Palestinian”, etc. or Africa (“Berber”, “Congolese”, “Zulu”, etc.).

Yardsticks

Here are some common yardsticks based on Census data. Not all are official or used by the Census Bureau. Some come from the Department of Housing and Urban Development or other federal agencies.

-- **Age dependency ratio**: The ratio of people who are young or old compared to those who are working age (typically 18 to 64). This is a yardstick of an area’s ability to support the needs of schools and the

elderly. It can be broken down into a **child dependency ratio** (ages 0 to 17 divided by ages 18 to 64) and an **old-age dependency ratio** (ages 65+ divided by ages 18 to 64.)

-- **Crowding:** A household with more than 1 person per room (bathrooms, kitchens, utility rooms and unfinished spaces don't count). **Severe crowding** is more than 1.5 people per room. Nationwide, that's 2.3% and 1%, respectively.

-- **Cost-burdened household:** Shelter costs eat up more than 30% of the household's income. This is typically the rent contribution standard for rental housing programs. In a **severely cost-burdened household**, shelter eats up more than 50% of the household's income. For owner-occupied units, the Census Bureau computes this based on **selected monthly owner costs**, which include mortgages, property taxes, insurance and utilities.

-- **Diversity:** How likely are two people chosen at random to vary by race and Hispanic ethnicity? One way to measure this is the USA TODAY Diversity Index. Invented in 1991 for just this purpose, it varied from 0 to 100 and can be applied to an area of any size.

-- **Headship rate:** This simply divides the number of adults (or any age group of them) by the number of that same group who are counted as householders. It's a measure of independence. During the Great Recession, the national rate for 18-to-34-year-olds dropped sharply, from 38% to 36%. Rates for older groups run as high as 65%. The higher the rate, the more independent the members of that age group are.

-- **Homeownership rate:** The share of households in an area whose occupants own it outright or hold a mortgage. The national rate has dropped several percentage points to the low 60s from record levels reached before the housing market meltdown and Great Recession.

-- **Low-income:** Typically defined as a household making 80% or less of the area's median household income. **Very low income:** Less than half of an area's median. **Extremely low income:** Less than 30% of an area's median.

-- **Linguistic isolation:** The Census Bureau recently abandoned this term under pressure from advocacy groups that called it insensitive, but you'll still see it in older data and documentation. It's applied to a household where no one over 14 speaks English at least "very well." (See below.) It's still an idea worth measuring. There's no simple replacement phrase.

-- **Naturally-occurring retirement community:** A neighborhood where at least 40% of householders are 65 or older.

-- **Segregation:** How different is the racial mix of a neighborhood compared with the city or the metro? Sociologists use this index to measure the "sorted-outness" of any two groups with respect to each other. It's sometimes explained as the percentage of the minority that would have to move around to distribute itself evenly across the larger area. Typically, it measures segregation of non-Hispanic whites compared with blacks or Hispanics, but it could be used to measure segregation by age, citizenship, income, education or other measures.

-- **Speaks English less than "very well":** Studies have shown that anyone who rates their English-speaking ability lower than "very well" (the other choices: "not at all" and "well") probably has trouble navigating daily life, such as reading a lease, an instruction manual, a bus schedule or a ballot.

-- **Unattached:** This term has gotten popular because of the Great Recession and the ailing male college attendance rate. It's used to measure the share of young adults – typically 16 to 25 – who are neither working or enrolled in school.