## Q&A Topics in Geographic Information Systems

## Thursday, July 11, 2013

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q: ­Do you know if there is an equivalent to American Fact Finder in Canada? ­

A: Per Kim Boyuk at Stats Canada “*­We do have something very similar in Canada. Statistics Canada is responsible for conducting the Census every five years. I’m including the link to Statistics Canada website and on the right sidebar you will see ‘Census of Canada’. ­ ­Click on that and it will lead you into a site similar to the American Fact Finder.* [*http://www.statcan.gc.ca*](http://www.statcan.gc.ca) *“*

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q: ­In the old version of Factfinder, we could search for census tract by specific address. Is that feature still available in AF2?­

A: ­Yes, on the main page there is an "Address Search" feature that provides this capability. Look for the same street sing icon as in previous versions.

For the latest screen shot, you can review slide 18. The feature is located on lower right and highlighted by a box.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q: ­How does census handle multiple races for individuals?­

A: The census reports race differently at different level of geographies. For the age by sex tables we are primarily interested in (at tract level), the race is reported as “WHITE ALONE”, “BLACK OR AFRICAN AMERICAN ALONE”, etc and then “SOME OTHER RACE ALONE” and “TWO OR MORE RACES”. This means that bridging based on factional assignment to match with the cancer data is difficult and only equal allocation can be used at this level.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q: ­How do you see the best way to map the poverty data along with cancer incidence of various cancers to help guide community efforts? I receive many requests to map cancer incidence and mortality along with various demographics. ­

A: It depends upon the research question. One basic descriptive approach is to overlay the cancer data with the SES data with hatchings—either a composite or separate maps for education, poverty, etc. If communities are interested in the association between cancer and area-based measure of SES, the rates by SES (as demonstrated in Dan’s section) would be useful but not mappable. Another approach is to do cluster detection, for instance in SaTScan (http://www.satscan.org/), to find areas of risk. These areas can then be described based on community profile. Or the influence can be adjusted for in the cluster model. A possibly useful resource for community mapping is the Community Health Needs Assessment (CHNA) website (http://assessment.communitycommons.org/CHNA/). They have an interactive “Health Indicators Map Gallery” feature with a large collection of data on area demographics, socio-economics, and health outcomes including some cancer data.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q: What state and US population do you use for age adjusted rates, for year of dx 2010?

A: To age-adjust rates, you use the Year 2000 Standard population. This standard does not change with each census.

For the underlying denominator, you could use the Census 2000 or Census 2010 SF1 data (whichever is closer to your dx years). However these are the populations for April 1 of the census year. The intercensal estimates (such as those available in SEER\*Stat) provide population estimates for the middle of each year (July 1) which is more appropriate for calculating annual rates. In this case, for cases dx year 2010, you would use the estimates for July 1, 2010.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q: The data Ferrett only include 5-year data, right?

A: No--DataFerrett also has single years of data. It actually has a whole host of interesting data including BRFSS data and Library data. The datasets are listed here: <http://dataferrett.census.gov/datasets.html> \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q: ­Can you please re-review how to specify 2000 vs 2010 census data?

A: They are labeled as such in American FactFinder. Using the advanced search, you select the 2000 or 2010 (or other) dataset by clicking on the “Topics” bar.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q: ­When will the data conversion take place from FIPS to ANSI?

A: The responsibility for maintaining the codes has been transferred to ANSI (American National Standards Institute). The codes remain the same and are still labeled FIPS codes. In the future, the responsibility for numbering new counties, countries, etc. will be the responsibility of ANSI instead of the National Institute of Standards and Technology (NIST).

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q: Thanks Dave, the AFF interface is finally getting easier to use. The first time I used it was just after they junked the old system. I wound up calling the census help desk after spending several hours on the site trying to find the P12 tables.

A: ­Yes - they junked the old system before the new system was really ready for prime time. At this point, it seems to be able to do everything the old system did and also has some nice new features.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q: ­Recinda, what is a good example of how cancer registries can use ACS data, rather than, for example, the P12 tables in the SF1 dataset?

A: The ACS data are our main source of social, economic, housing, and linguistic data. For cancer control, we may be interested in screening a community, so we could use ACS data to identify if materials are needed in non-English languages or how many are uninsured. Another example is we know poverty influences health, so we can both describe rates based on poverty as well as adjust for poverty or education in our models of risk. ACS data can also be used when there are community concerns about higher cancer rates to provide insight on whether the high rates might be based on an exposure or demographic profile.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

 Q: ­What are some of reasons of over and under estimates?

A: There are two main reasons. The estimates are based on births and deaths and migration. Migration is the least accurate although they have drastically improved collection methods, so the end of decade issues in 2010 were less than 2000. But, for instance, folks don't fill out tax returns so we don't know when they move—this impacts the elderly and the poor in particular­. These collection/tracking issues can cause problems at all levels of geography.

The second reason is that the county level data are estimated (a type of fractional allocation) based on state level. Data on migration between counties is not as good as migration between states.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q: ­What data should we be using for denominators when we're interested in smaller areas, i.e. census tracts?

A: Denominator should come from the decennial census i.e. SF1/2010. It will get stale, which is why you would use intercensal estimates--but usually they are not released below county. You can create your own using fractional allocation (as is demonstrated in the bridging section) to proportionally populate tract data based on county data. But of course, this can get complicated and introduce error.