Creating Your Own Census Tract-Level SEER\*Stat Databases NAACCR Version 22

Q&A

January 30, 2023

|  |  |  |
| --- | --- | --- |
| **#** | **Question** | **Answer** |
| 1 | Can we do this for the entire US? | Yes, the populations are available for the entire US. |
| 2 | Do you exclude all records without Census tract information? | No records are excluded, even if they have missing or incorrect census tract 2010 values. These records will of course not contribute to statistics by census tract, but they can be used in county or state analyses. They will also be available so you can characterize the quality of your geocoded data. |
| 3 | Has anybody created process this for R, or only SAS? | We have not developed a process in R. If you can read the XML with R, then doing the rest would not be difficult. |
| 4 | Does the program calculate age-adjusted incidence rates at the census tract level? | The SAS program is used to prepare data for SEER\*Prep. SEER\*Prep creates databases to be analyzed using SEER\*Stat. SEER\*Stat can be used to calculate crude and age-adjusted rates, and other statistics, at the census tract level. |
| 5 | Sorry, I didn't quite understand what the section of code creating the combinations by state/county/tract/year for the incidence and population files - what was contained in the output for that piece of code? | It does 2 things: 1. It will add to the attribute data file (for input into SEER\*Prep), any state-county-tract year combinations that are in the incidence data or population data that were not already in the attributes (primarily this is for unknown/missing county or tracts) 2. It will create a file with every combination of state-county-tract in your data, to use as a format in SEER\*Prep. This was demonstrated during the portion of the webinar that created the attributes database in SEER\*Prep. |
| 6 | This looks like fixed-length format, not csv | That is correct. Currently only the incidence datasets are in CSV format - all other data (populations, mortality, attributes, etc.) are still fixed width, but that will change in the coming months. And then we will eventually add support for NAACCR XML formatted incidence files. |
| 7 | Can the created census-tract attributes be used in caselisting? | Yes |
| 8 | Would you use a user defined variable for state level legislative districts? | If your state-level legislative districts were built using census tract 2010 geography, you could use either a user-defined variable or you could add the field to the attributes database. If your state-level legislative districts were built using a lower level geography than tract, like block, then this will not work. |
| 9 | Do you want to talk a bit about ses-stratified survival? | You can calculate survival statistics stratified by any of the variables in the linked census tract attributes database. However, the life tables used for relative survival are the same ones used for national, state, and county statistics. We have not built census tract-level life tables. Thus, it may be more appropriate to analyze cause-specific survival by census tract-based SES than relative survival. |