

Joining data assignment

In this assignment, you're going to join two datasets together: Small Area Health Insurance Estimates data from two different years:

- [A SQLite database that has the data already imported into it.](#)
- [The documentation for the Small Area Health Insurance Estimates \(sahie2013\) data.](#)
- [The Gentle Introduction to SQL tutorial](#)

For purposes of this assignment, you are a data analyst who has been asked to evaluate the effect of Obamacare on the percentage of uninsured people in the United States. To do this you will need to, in SQL, do the following:

- Select the fields you need.
- Join the `sahie2010` and `sahie2015` tables based on the `GEOID` field.
- Filter the results to show the widest range of people possible. There should be just one number per county inclusive of the most people possible. Look at the documentation for help on this. You will be writing `WHERE ... AND ...` statements.
- Filter out statewide totals. In the database, `geocat "40"` is statewide totals. `Geocat "50"` is county.
- Calculate the percent change between the two years.

NOTE: While it is possible to do all of this in a single line of SQL, the cautious analyst breaks things down into steps.

To get credit for this assignment, you must submit your SQL query or queries. I am going to run it on my copy of the database and it should just work.