

MABLE/GEOCORR Read-me file

Written by:

Andrew Foote
PhD Candidate
Department of Economics
University of California, Davis

afoote[at]ucdavis[dot]com
andrewfoote.weebly.com

Updated April 7, 2014

MABLE/GEOCORR LINK

2010 GEOGRAPHY: <http://mcdc.missouri.edu/websas/geocorr12.html>

2000 GEOGRAPHY: <http://mcdc.missouri.edu/websas/geocorr2k.html>

1990 GEOGRAPHY: <http://mcdc.missouri.edu/websas/geocorr90.shtml>

PURPOSE

The purpose of this users guide is to outline the basic use of the MABLE/GEOCORR Engine, which is linked below above.

DISCLAIMER

This read-me file is neither official nor endorsed by the Missouri Census Data Center, which created this data product. It is simply a quick user's guide that I wrote in order to help fellow researchers, and to save me time so that I could refer people to my page rather than explain it multiple times.

Any errors are my own. I welcome suggestions on ways to improve this short user's guide.

INPUTS FOR MABLE/GEOCORR

In what follows, I outline the steps for how to create a crosswalk. Say you have data at the geography level of X, but you want to match X to Y, where Y is a larger geographic unit. I will use

counties and MSAs as the example, since it is more straightforward. So you have counties in your data, but you need to figure out which MSA they belong to.

1. STATES – Select the states for which you want to have a crosswalk (e.g. If you want all counties in MSAs in Virginia, just select Virginia in this step); you should probably select all here, just to be sure.
2. SOURCE GEOCODE – The geographic level that you already have; in my example, you would select county. There are a number of other options, such as census tract and census block.
3. TARGET GEOCODE – The geographic unit you want to crosswalk to. In this example, you would select MSA (actually PMSA). Note that in the 2010 version, PMSA is not an option; for this, you should use the 2000 version.
4. WEIGHTING VARIABLE – This is only useful if the crosswalk is not a one-to-one mapping. For instance, if you have ZIP codes to counties, there are some zipcodes that have area in two (or more) counties. For a good example, ZIP code 95616 (the zipcode for Davis, CA) actually falls into two different counties: Yolo and Solano. The question then is how to take this into consideration. There are two main ways to weight:
 - a. Population – useful if you are looking at employment or something residential. Since the only part of 95616 that is in Solano is very sparsely populated, it would give a low weight to that (the variable AFACT).
 - b. Land Area – useful more for temperature readings, or anything that where actual space is important. I don't generally use this weight because it doesn't apply to much of my research.
5. Then hit “run request” – it will generate as CSV that you can save by right-clicking. The resulting file will have three main variables: the source code, target code, and the allocation factor (afact) which is based on the weighting variable you choose in step 4. If the source code has only one target code that matches it, $afact = 1$. If the source code has multiple target codes, $afact$ will be between 0 and 1. In practice, most values of $afact$ are either close to zero or close to 1.

Short note: Practical issues with the weighting variable

I want to spend a little time talking about how one might address the fact that the allocation factor may not be 1. Let's take an example of ZIP codes as the source, counties as the target, and population as weighting factor.

Below are three ZIP codes in Missouri that are in a few different counties.

ZIP CODE	COUNTY FIPS	COUNTY NAME	ZIP NAME	POP 2010	AFACT
63014	29071	Franklin MO	Berger, MO	688	0.883
63014	29073	Gasconade MO	Berger, MO	91	0.117
63015	29071	Franklin MO	Catawissa, MO	1631	0.786
63015	29099	Jefferson MO	Catawissa, MO	445	0.214
63016	29099	Jefferson MO	Cedar Hill, MO	7582	1
63017	29189	St. Louis MO	Chesterfield, MO	41470	1
63019	29099	Jefferson MO	Crystal City, MO	4449	1

63020	29099	Jefferson MO	De Soto, MO	20667	0.998
63020	29221	Washington MO	De Soto, MO	47	0.002

In the above table, there are three zip codes that are entirely contained in a single county – 63016, 63017, and 63019. However, three other zip codes, 63014, 63015, and 63020 are in two counties each.

For 63014, 88% of its population lies in Franklin County, MO, while the remaining 12% lies in Gasconade County, MO. 63015 is more split: 79% to 21%. 63020 is barely split: effectively 100% of its population lives in Jefferson County, MO.

How should a researcher proceed? It depends on the situation. If you have employment numbers at the zip code level, but want to do county-level comparisons, you might want to allocate the employment numbers to counties based on the allocation factor. However, researchers could do this in combination with a simple threshold rule, where a zip code is assigned to only one county if it has more than 90% of the population (or land area). However, researchers ought to test the importance of these thresholds, and their impact on final results.