

[Page One](#)
[Campus  
Computing  
News](#)
[IT Help is Here](#)
[Information  
Security  
Training Now  
Available  
Through WebCT  
Vista](#)
[UNT Web sites  
and Web  
Applications  
Need to be  
Registered by  
May 31](#)
[Summer Hours](#)
[Bits -n- Bytes](#)
[Today's  
Cartoon](#)
[RSS Matters](#)
[The Network  
Connection](#)
[Link of the  
Month](#)
[Helpdesk FYI](#)
[Short Courses](#)
[IRC News](#)
[Staff Activities](#)
[Subscribe to  
Benchmarks  
Online](#)

# Research and Statistical Support

## University of North Texas

### RSS Matters

Link to the last RSS article here: [Mapping And Data Representation In Stata: Part 1](#) - Ed.

## Mapping And Data Representation In Stata: Part 2

By [Patrick McLeod](#), Research and Statistical Support Services Consultant

In April's RSS Matters column, we discussed the basics of mapping data into a graph in Stata. Now that we have all the pieces in place, in Part 2 we will use Stata and the Stata packages discussed in Part 1 to produce a choropleth map of the United States (the Lower 48 plus Alaska and Hawaii) showing the murder rate in 2003 by state.

The first step in this process is to unzip the Shapefile or MapInfo file that you are using to render the map as a graph in Stata. For this example, we are using the United States states and territories Shapefile downloaded from the National Weather Services' website. Once this file has been unzipped to my Mac desktop, I issue the following command:

```
shp2dta using /Users/patrick/Desktop/s_01au07/s_01au07, database(usdb) coordinates(uscoord)
genid(id)
```

The Stata package shp2dta renders the Shapefile as a Stata data file. In this case, the Stata data file is named *usdb* and the very important state ID variable *id* is generated based on the state coordinates from the *uscoord* file.

From this point on in the process, producing your map in Stata is a matter of working with Stata datasets (doing some data editing and using the *merge* command) and producing the graph you want using Stata's powerful suite of graphing commands.

```
use "/Users/patrick/data/state_rates_2003"
```

Now that you have your source data (the data that you want to map) and your Stata data file from the Shapefile for the map, there are two data management steps to be aware of before we move forward with more Stata code: Making sure the *id* variable from the map file has the correct states in it and that it matches up to your state names in your source data. When you download the file from the National Weather Service website, unzip it and enter it into Stata using the shp2dta command, you will notice that this file renders all 50 states, the District of Columbia and all U.S. territories (American Samoa, Guam, Midway, Puerto Rico and the U.S. Virgin Islands).

Most of the time when you are dealing with social science data that is explicitly state-based, you

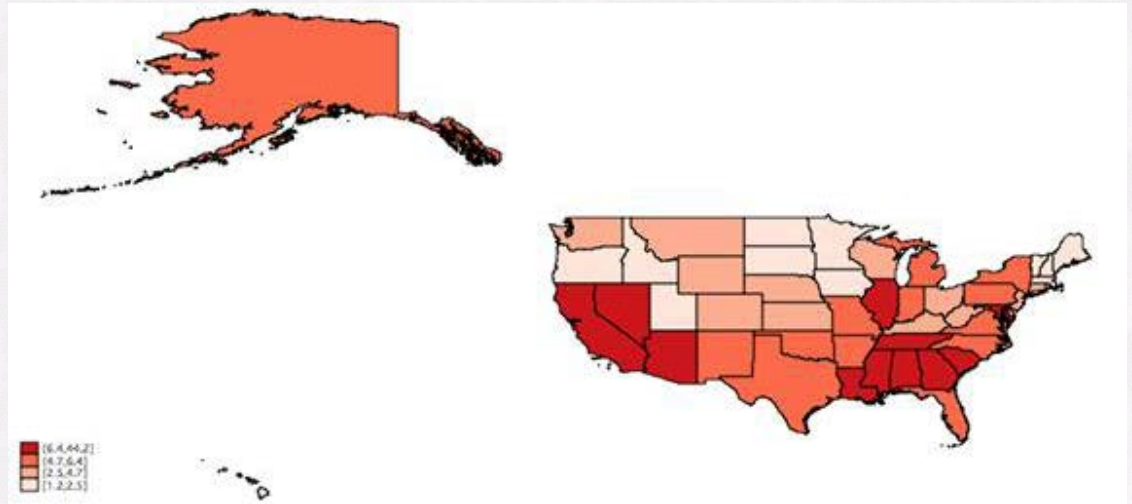
won't have a use for the territories that are included in the Shapefile. My suggestion is to simply manually drop them before you try to merge your data so that you are left with a map data file that has 50 states plus D.C., the structure in which most of your data will be organized. While it might not be the best way to do it, the most direct way to do this is to sort your source data file by state name and save it, then sort your map file by state name and save it. Copy the *id* variable from the map data file and paste it into the source data file. Now both of your data files are sorted on the same *id* variable and are ready to merge on that same variable:

```
merge id using "/Users/patrick/Desktop/s_01au07/usdb"
```

After you successfully issue this command, you now have all the components in one data file to produce a map: Your *id* variable, your source data and your map data. To produce your map in Stata, issue the following command:

```
spmap murderrate using uscoord, id(id) fcolor(Reds)
```

This generates the map below (exported in .PNG format and cropped in PhotoShop):



You could use many of the Stata commands for formatting graphs to spice up your final product, for instance including the option *b1title*(*Murder Rate By State, 2003*) to add a title to the graph so that it will have a title printed at the bottom that reads 'Murder Rate By State, 2003'.

Until next time, happy computing and best of luck with your choropleth mapping!

Originally published, May 2008 -- Please note that information published in *Benchmarks Online* is likely to degrade over time, especially links to various Websites. To make sure you have the most current information on a specific topic, it may be best to search the UNT Website - <http://www.unt.edu> . You can also search *Benchmarks Online* - <http://www.unt.edu/benchmarks/archives/back.htm> as well as consult the UNT Helpdesk - <http://www.unt.edu/helpdesk/> Questions and comments should be directed to [benchmarks@unt.edu](mailto:benchmarks@unt.edu)

[Return to top](#)

