

[Page One](#)
[Campus Computing News](#)
[Photoshop Express - Free Online](#)
['Photoshop' is a pretty good deal](#)
[Computer-Based Training for Microsoft Project 2007 is Now Available](#)
[Reports of Windows Collapsing are Overblown but There IS a Lesson to be Learned Here](#)
[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: [New Developments With R - version 2.6.2 - Ed.](#)

Mapping And Data Representation In Stata: Part 1

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

Dealing with geographic components in social science research can take many forms. Data often have some geographic component to them (the distribution of responses to a survey or the location of nations suffering through civil war, for example). Accounting for these geographic components often takes the form of dummy coding variables or employing techniques such as geographically weighted regression. The most commonly encountered geographic component is in the display of data or the display of results. Whether your research deals in national level data, U.S. state level data, Congressional district level data or any other data that can be displayed within boundaries, maps offer an excellent way to display spatial data that are organized in ways that are visually compelling and easy for readers to comprehend.

Stata offers a user contributed package called *spmap* to generate maps with data as graphs. *spmap* will work with both Stata 9 and Stata 10. In order to use *spmap*, you'll need to install *spmap*, *shp2dta* and *mif2dta* using Stata's internet functionality. If this is your first time to install something via the internet from within Stata, the easiest thing to do is to type the following three commands into the command line:

```
findit spmap
```

```
findit shp2dta
```

```
findit mif2dta
```

and then click on the blue text items (blue text means an internet linkage, exactly like an embedded web link on a web page) to download and install these three packages within your Stata installation.

From the help file for *spmap*, here's an overview of the package:

spmap is aimed at visualizing several kinds of spatial data, and is

particularly suited for drawing thematic maps and displaying the results of spatial data analyses.

spmap functioning rests on three basic principles:

- First, a base map representing a given region of interest *R* made up of *N* polygons is drawn.
- Second, at the user's choice, one or more kinds of additional spatial objects may be superimposed onto the base map. In the current version of *spmap*, six different kinds of spatial objects can be superimposed onto the base map: polygons (via option `polygon()`), polylines (via option `line()`), points (via option `point()`), diagrams (via option `diagram()`), arrows (via option `arrow()`), and labels (via option `label()`).
- Third, at the user's choice, one or more additional map elements may be added, such as a scale bar (via option `scalebar()`), a title, a subtitle, a note, and a caption (via `title_options`).

Proper specification of *spmap* options and suboptions, combined with the availability of properly formatted spatial data, allows the user to draw several kinds of maps, including choropleth maps, proportional symbol maps, pin maps, pie chart maps, and noncontiguous area cartograms.

While providing sensible defaults for most options and suboptions, *spmap* gives the user full control over the formatting of almost every map element, thus allowing the production of highly customized maps.

The help file for *spmap* is extensive; instead of replicating the whole file in this article, I suggest that as soon as you finish reading this article you install *spmap* and then view and/or print the help file.

The *spmap* package allows users to create [choropleth maps](#) (link opens a file in PDF). A choropleth map displays data values or phenomena across different regions in shades of color or different gradients. The example we're going to work with in this article will map homicides per capita onto a map of the United States with different levels of homicides per capita displayed in different shades of red; lighter red reflects lower homicide levels while darker red reflects higher homicide levels.

Before we can get down to brass tacks in Stata, however, we need find a boundary file (a map) that is in either `.shp` (shapefile) or `.mif` (MapInfo interchange format). The best way to find one of these files is via the web; for instance, Googling "United States" shapefile or "United States shapefile" returns loads of possibilities. Per the [Stata FAQ page for working with spmap](#), I'm going to use the U.S. shapefile from the [National Weather Service's website](#).

Shapefiles come from ESRI, Environmental Systems Research Institute, one of the leading firms in the field of GIS software and development. ESRI produces the widely-used [ArcGIS](#) software platform which can be found on some computers here at UNT. A shapefile is an ESRI-specific map. Shapefiles are generally easier to find than MapInfo format files. The Stata package *shp2dta* converts polygon shapefiles into Stata data format. Note the restriction to polygon shapefiles!

MapInfo interchange format files work with the [MapInfo](#) suite of software products. In my experience, *.mif files are not as common as shapefiles but are worth looking for if your shapefile options are limited or if you cannot find a shapefile for your particular needs. The Stata package `mif2dta` converts a *.mif file into a Stata dataset for processing.

At this point in the process of generating your map as a graph in Stata, you should have the following things: `spmap`, `shp2dta`, `mif2dta`, a shapefile format map of your intended Stata map or a MapInfo format map of your intended Stata map. In Part 2 of the article we will walk through the Stata code involved in mapping data to a map and producing a graph figure of that map.

Originally published, April 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

[Return to top](#)