

**spshape2dta** — Translate shapefile to Stata format

[Description](#)  
[Options](#)
[Quick start](#)  
[Remarks and examples](#)
[Menu](#)  
[Also see](#)
[Syntax](#)

## Description

`spshape2dta` *name* reads files *name.shp* and *name.dbf* and creates Sp dataset *name.dta* and translated shapefile *name\_shp.dta*. The translated shapefile will be linked to the Sp dataset *name.dta*.

## Quick start

Create *myfile.dta* and *myfile\_shp.dta* from *myfile.shp* and *myfile.dbf*

```
spshape2dta myfile
```

Create *newfile.dta* and *newfile\_shp.dta* from *oldfile.shp* and *oldfile.dbf*

```
spshape2dta oldfile, saving(newfile)
```

## Menu

Statistics > Spatial autoregressive models

## Syntax

```
spshape2dta name [ , options ]
```

<i>options</i>	Description
<code>clear</code>	clear existing data from memory
<code>replace</code>	if <i>name.dta</i> or <i>name_shp.dta</i> exists, replace them
<code>saving(<i>name2</i>)</code>	create new files named <i>name2.dta</i> and <i>name2_shp.dta</i> instead of <i>name.dta</i> and <i>name_shp.dta</i>

`spshape2dta` translates files *name.shp* and *name.dbf*. They must be in the current directory.

`spshape2dta` creates files *name.dta* and *name\_shp.dta*. They will be created in the current directory. The data in memory, if any, remain unchanged.

## Options

`clear` specifies to clear any data in memory.

`replace` specifies that if the new files being created already exist on disk, they can be replaced.

`saving(name2)` specifies that rather than the new files being named `name.dta` and `name_shp.dta`, they be named `name2.dta` and `name2_shp.dta`.

## Remarks and examples

[stata.com](http://www.stata.com)

`spshape2dta` is the first step in preparing data to be used with shapefiles. See [\[SP\] Intro 4](#) for step-by-step instructions.

`spshape2dta` creates two files:

```
name.dta
name_shp.dta
```

`name.dta` is an ordinary Stata dataset. The dataset will have  $N$  observations, one for each spatial unit. The dataset will be `spset`.

```
. use name
. spset
  Sp dataset
           data:  cross sectional
  spatial-unit ID:  _ID
  coordinates:  _CY, _CX (latitude-and-longitude, miles)
  linked shapefile:  name_shp.dta
```

`name.dta` will contain the variables

```
_ID          values 1, 2, ..., N. This variable links observations in the
              data to observations in the Stata-format shapefile, name_shp.dta.
_CX, _CY     contain the centroids for the places (spatial units)
```

`name.dta` will include the other variables defined in `name.dbf`. Usually, there will be five or ten. What they contain varies but can usually be determined from their names and by looking at their values.

`name.dta` will be linked to `name_shp.dta`, which is called the Stata-format shapefile. It contains the map. It too is an ordinary Stata dataset, but you ignore it. `Sp` will use `name_shp.dta` behind the scenes when you construct contiguity spatial weighting matrices using `spmatrix create contiguity` or when you graph choropleth maps using `grmap`.

## Also see

[\[SP\] Intro 3](#) — Preparing data for analysis

[\[SP\] Intro 4](#) — Preparing data: Data with shapefiles