Title

> **spdistance —** Calculator for distance between places

Description

spdistance $\#_1$ $\#_2$ reports the distance between the areas $\_\text{ID} = \#_1$ and $\_\text{ID} = \#_2$.

collect is allowed; see [U] **11.1.10 Prefix commands**.

Quick start

Obtain distance between $\_\text{ID} = 48201$ and $\_\text{ID} = 48041$

    spdistance 48201 48041

Menu

Statistics > Spatial autoregressive models

Syntax

    spdistance $\#_1$ $\#_2$

$\#_1$ and $\#_2$ are two $\_\text{ID}$ values.

Remarks and examples

Remarks are presented under the following headings:

> *Are coordinates really planar and not latitude and longitude?*
> *Reverse engineering planar distances*
> *More than you want to know about coordinates*
>> *Planar coordinates*
>> *Latitude and longitude coordinates*

Are coordinates really planar and not latitude and longitude?

The purpose of spdistance is to help in understanding the units in which distances are measured when coordinates are recorded in planar units. Before turning to that issue, however, let us ask another question: Are the coordinates recorded in your data really planar? Sp assumes that they are. It is your responsibility to change a setting if they are in fact degrees latitude and longitude. You change the setting by typing

    . spset, modify coordsys(latlong, kilometers)

or

```
. spset, modify coordsys(latlong, miles)
```

If coordinates are latitude and longitude, type one of those commands and then distances will be reported in kilometers or miles and you can dispense with spdistance for determining units.

Why does Sp assume that coordinates are planar when they might be latitude and longitude? How can you tell whether your data contain latitudes and longitudes?

Sp assumes that coordinates are planar because coordinates obtained from shapefiles are supposed to be planar. Usage is running ahead of standards, however, and these days many shapefile providers are providing latitude and longitude.

Before answering the second question, let us answer a third question you may be asking yourself: "Do I care? How bad would it be to treat degrees as if they were planar?" If all the locations in your data are near the equator, there is no reason you should care. But this is not likely to be the case, and because degrees longitude are not a fixed distance, your calculations will be incorrect if you treat degrees as if they were planar. See *Latitude and longitude coordinates* below for more details.

So how do you tell the units of measure? The documentation for your shapefile may tell you. If not, you can inspect the data. Sp datasets record the coordinates in variables _CX and _CY. You look at those variables and compare them with latitudes and longitudes for the same places or nearby places, which you can easily find on the web. If coordinates are latitude and longitude, then

- _CX will be the longitude value
- _CY will be the latitude value

Reverse engineering planar distances

Planar coordinates have no predetermined scale. Two places might be 5 apart. How far is that? One way to find out is to reverse engineer the scale. Take two places that you know the distance between, use spdistance to obtain the planar distance, and divide.

For instance, we have a city dataset in which Los Angeles and New York have _ID values 1 and 79. Using spdistance, we obtain the distance between them.

```
. spdistance 1 79
  (data currently use planar coordinates)
```

_ID	(x, y) (planar)
1	(0, 0)
79	(4.97, 1)

distance	5.0696053 planar units

The distance between the cities is roughly 2,400 miles, so we know that one planar unit equals $2400/5.07 \approx 473$ miles.

# More than you want to know about coordinates

Planar coordinates

Planar coordinates, also known as rectangular coordinates, are coordinates on a plane. The formula for the distance between two places (x_1, y_1) and (x_2, y_2) is given by the Euclidean distance formula:

$$\text{distance} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Distance will be measured in miles if x and y are measured in miles.

Latitude and longitude coordinates

Latitude and longitude are measured in degrees on a globe. The common illustration looks like this:
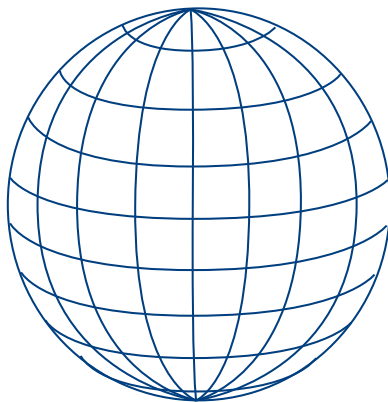
The vertical lines passing through the poles are lines of equal longitude. Longitude indicates which vertical line you are on. the vertical lines. It measures which vertical line you are on. Longitude is an east–west measure. Examples of longitude include $150°$ east and $96°$ west, with east and west referring to east and west of Greenwich, UK.

The horizontal rings are lines of equal latitude. Different horizontal lines are different latitudes. Latitude is a north–south measure. Examples of latitude include $30°$ north and $33°$ south, with north and south referring to north and south of the equator.

College Station, USA, and Sydney, Australia, are located at

City	Latitude	Longitude
College Station	$30°36'05''$ N	$96°18'52''$ W
Sydney, Australia	$33°51'54''$ S	$151°12'34''$ E

Computers use signed values to indicate direction and fractions of degrees instead of minutes and seconds. The above table can equivalently be written as

City	Latitude	Longitude
College Station	30.601	−96.314
Sydney, Australia	−33.865	151.209

If $1°$ equaled a fixed distance, we could use the Euclidean formula for calculating the distance between College Station and Sydney.

One degree of latitude does equal a fixed distance, namely, 69 miles, if the Earth were a sphere.

One degree of longitude, however, measures a distance that varies from 69 miles at the equator to 0 miles at the North and South Poles:

At latitude	$1°$ longitude equals	
±0	69 miles	(equator)
±10	68	
±20	65	
±30	60	
±40	53	
±50	44	
±60	35	
±70	24	
±80	12	
±89	1	
±90	0	(N or S pole)

There are formulas for calculating distances from latitude and longitude, and Sp will use them if you tell it that the coordinates are degrees latitude and longitude. If you do not, Sp makes calculations using the Euclidean formula, and that will result in incorrect distances except at the equator. At the equator, $1°$ longitude equals $1°$ latitude equals 69 miles. The farther north or south you make the calculation, the more will be the error. East–west distances will be exaggerated relative to north–south distances, resulting in places being calculated as being farther apart than they really are.

The 48 contiguous states of the United States lies between $25°$ and $50°$ latitude, a region in which $1°$ longitude varies between 66 and 44 miles.

Europe lies between $35°$ and $70°$ latitude, a region in which $1°$ longitude varies between 56 and 24 miles.

Stored results

spdistance stores the following in r():

Scalars
    r(dist)   distance between

Macros
    r(coordsys)   planar or latlong
    r(dunits)   miles or kilometers if r(coordsys) = latlong

Methods and formulas

If coordinates are planar, the distance between (x_1, y_1) and (x_2, y_2) is

$$d = \sqrt{(x_2 - x_1)^2 + (y_1 - y_1)^2}$$

If coordinates are latitude and longitude, let (t_1, n_1) and (t_2, n_2) be the two coordinate pairs, where t represents latitude and n represents longitude converted from degrees to radians.

Let $\Delta t = (t_2 - t_1)$ and $\Delta n = (n_2 - n_1)$. Then the distance between the two points is

$$d = r \text{ invhav}\{\text{hav}(\Delta t) + \cos t_1 \cos t_2 \text{ hav}(\Delta n)\}$$

where r is the radius of the Earth measured in the desired units (miles or kilometers) and

$$\text{hav}(\theta) = \frac{1 - \cos \theta}{2}$$

$$\text{invhav}(h) = 2 \text{ asin}(\sqrt{h})$$

Reference

Weber, S., and M. Péclat. 2017. A simple command to calculate travel distance and travel time. *Stata Journal* 17: 962–971.

Also see