An Alternative to GIS: Presenting Data on Maps Using Stata

George Woolsey
San Diego Gas & Electric Co.
WLRA Spring 2012 Conference
10 October 2012

Overview

- The common perception is that the only tool capable of working with maps and geographic information is Ersi's ArcGIS
- Most software programs that load researchers already use have these capabilities
  - Stata: `geocode`, `spmap`, `shp2dta`, etc.
  - SAS: MAPIMPORT, GMAP, etc.
  - R: maps, mapproj, etc.
- Circumvent the GIS department and save money on consultant fees and additional software

Maps in Stata

The commands:

`shp2dta`

Converting shape boundary files (shapefiles) to format usable by Stata

`spmap`

Visualizing several kinds of spatial data

`geocode`

Geocode addresses and calculate latitude and longitude

`vincenty`

Calculating distances between a pair of points

Example:

I have data by zip code:

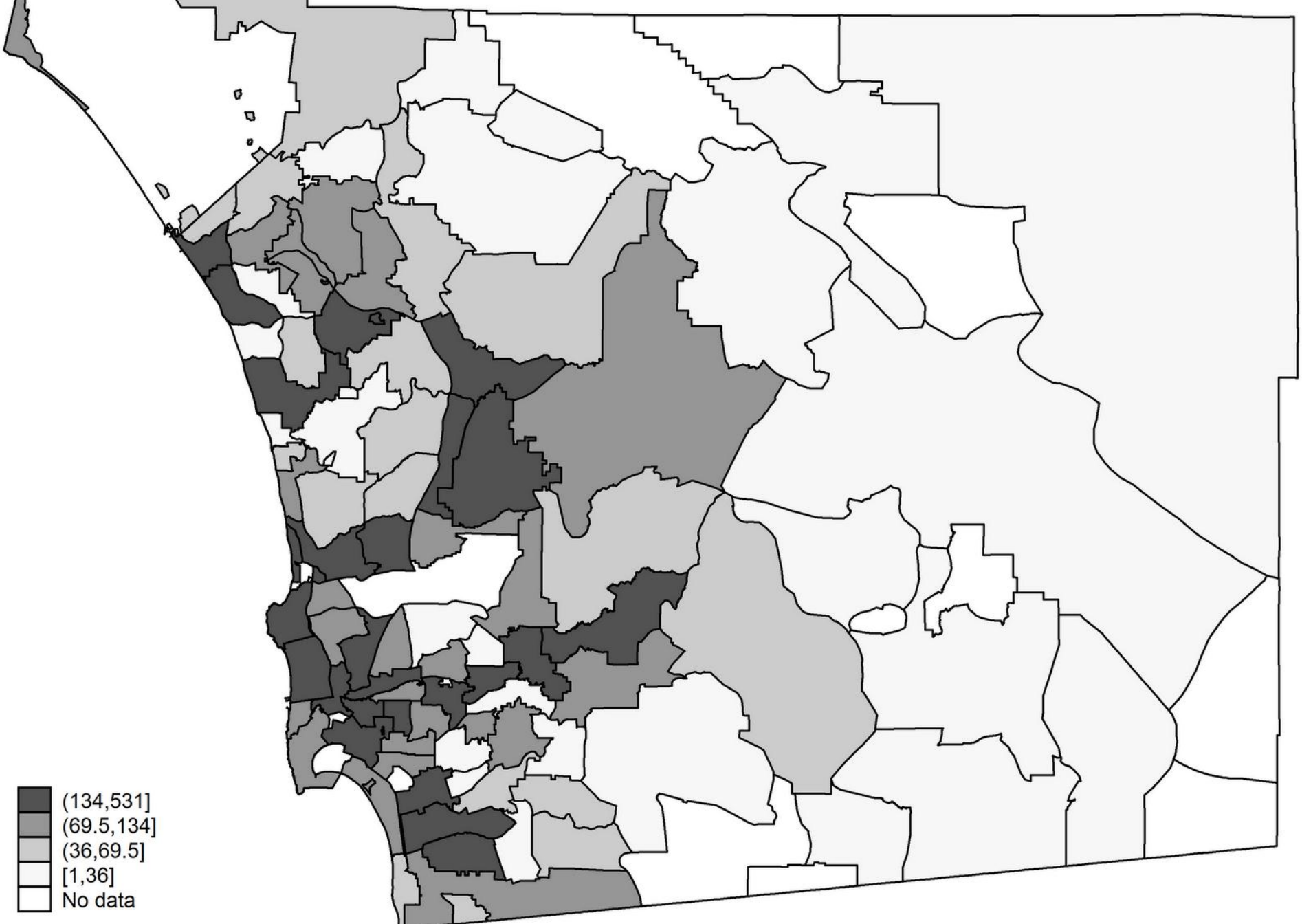
Zip Code	Restaurants	BK	McDs	SWay
91901	50	0	0	3
91902	39	0	1	1
91905	3	0	0	0
91910	191	3	2	5
91911	151	1	2	2

1. Search the web for a map: "San Diego County zip code shapefile"
2. Use `shp2dta` command to convert the shapefile data (.shp & .dbf) to format usable by Stata.

```
shp2dta using Zip_Codes.shp, datbase(mydbase) coordinates(mycoords)
```

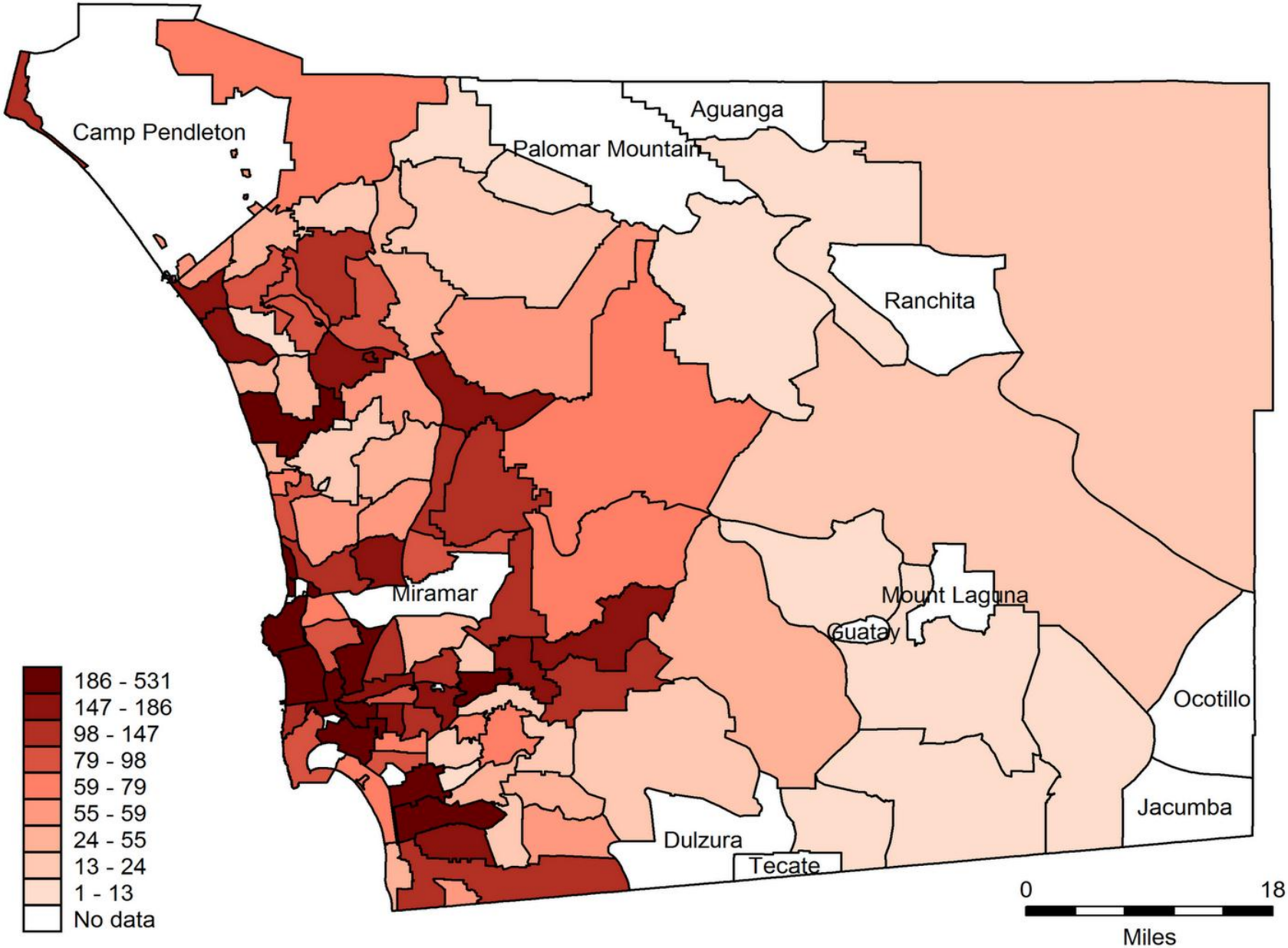
3. Merge my file with the "`mydbase`" file created by the `shp2dta` command based on zip code

```
spmap Restaurants using "mycoords.dta", id(_ID)
```



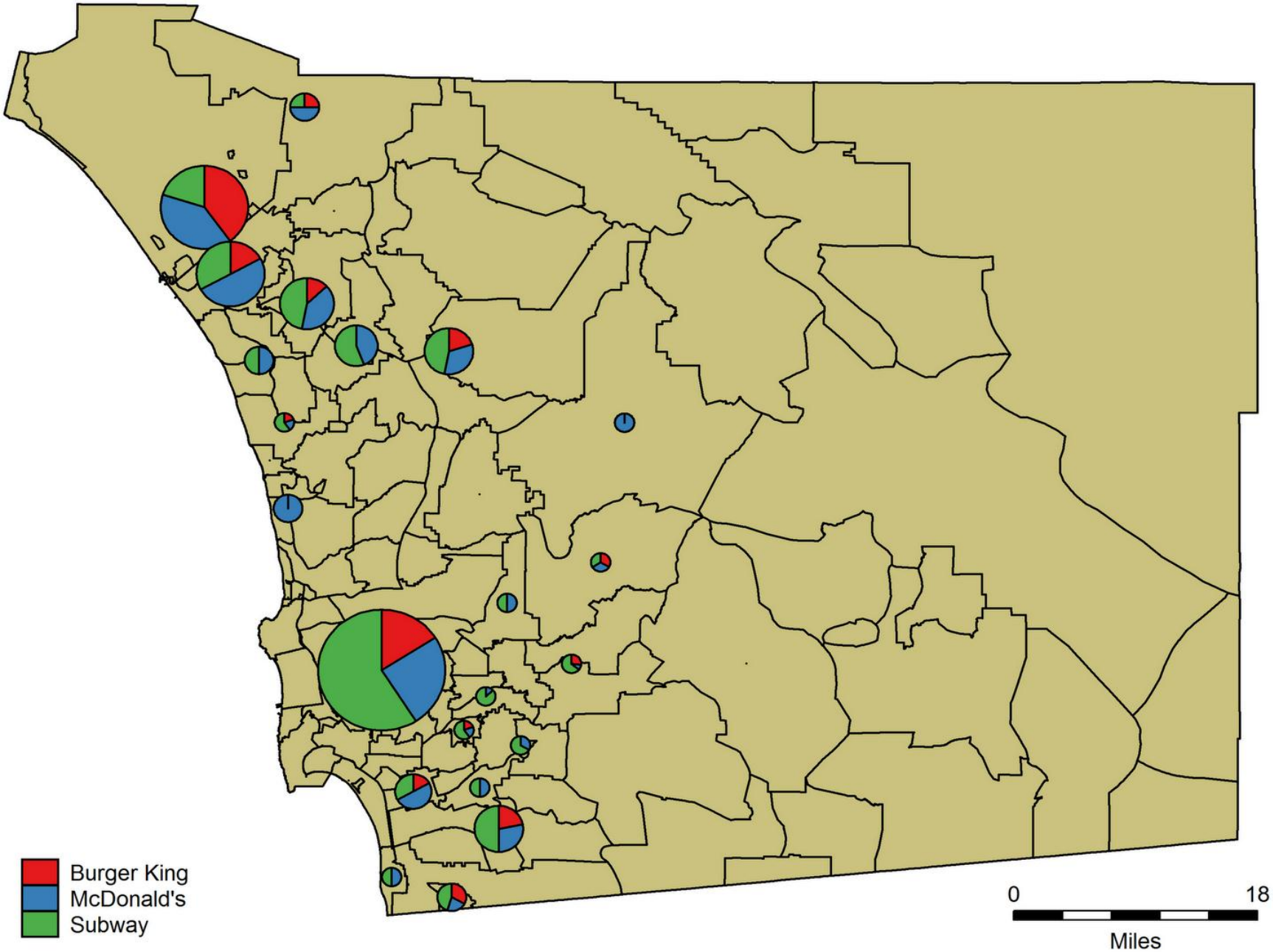
San Diego County Restaurant Locations

by Zip Code



Note: Communities with no restaurant data available are labeled. Prepared by G.Woolsey 5Oct2012

San Diego Fast Food Locations



Note: Plot size proportional to number of McDonald's

Geocoding

ID	Restaurant	Address	City	State	Zip	Latitude	Longitude
31	Subway	8055 ARMOUR ST	SAN DIEGO	CA	92111		
32	Subway	5523 MISSION RD	BONSALL	CA	92003	33.29076	-117.2262
33	Subway	950 E VISTA WAY	VISTA	CA	92084		
34	Subway	137 S LAS POSAS RD	SAN MARCOS	CA	92078	33.141597	-117.19051

```
geocode, address(Address) city(City) state(State) zip(Zip)
```

ID	Restaurant	Address	City	State	Zip	Latitude	Longitude	Geocode	Geoscore
31	Subway	8055 ARMOUR ST	SAN DIEGO	CA	92111	32.81966	-117.15	200	8
33	Subway	950 E VISTA WAY	VISTA	CA	92084	33.21142	-117.2308	200	8

geocode definitions

200 = No errors, successful parse and geocode
400 = Bad address request. The address is incorrectly specified
500 = Bad server request. Unknown failure reason.
601 = Empty Address
602 = Unknown address. May indicate a new or incorrect address.
603 = Address cannot be geocoded due to legal or contractual reasons
620 = Too many queries in 24 hour period

geoscore definitions

0 = Unknown accuracy.
1 = Country level accuracy.
2 = Region (state, province, prefecture, etc.) level accuracy.
3 = Subregion (county, municipality, etc.) level accuracy.
4 = Town (city, village) level accuracy.
5 = Post code (zip code) level accuracy.
6 = Street level accuracy.
7 = Intersection level accuracy.
8 = Address level accuracy.
9 = Premise (building name, property name, shopping center, etc.) level accuracy.

Calculating Distance

ID	Restaurant	Address	City	State	Zip	Latitude	Longitude	MyLat	MyLong
31	Subway	8055 ARMOUR ST	SAN DIEGO	CA	92111	32.81966	-117.15	32.978455	-117.25003
33	Subway	950 E VISTA WAY	VISTA	CA	92084	33.21142	-117.2308	32.978455	-117.25003
32	Subway	5523 MISSION RD	BONSALL	CA	92003	33.29076	-117.2262	32.978455	-117.25003
34	Subway	137 S LAS POSAS RD	SAN MARCOS	CA	92078	33.141597	-117.19051	32.978455	-117.25003

```
vincenty Latitude Longitude MyLat MyLong, hav(Haversine_Distance)
```

ID	Restaurant	Address	City	State	Zip	Latitude	Longitude	MyLat	MyLong	Haversine Distance
31	Subway	8055 ARMOUR ST	SAN DIEGO	CA	92111	32.81966	-117.15	32.978455	-117.25003	12.405657
33	Subway	950 E VISTA WAY	VISTA	CA	92084	33.21142	-117.2308	32.978455	-117.25003	16.125516
32	Subway	5523 MISSION RD	BONSALL	CA	92003	33.29076	-117.2262	32.978455	-117.25003	21.610093
34	Subway	137 S LAS POSAS RD	SAN MARCOS	CA	92078	33.141597	-117.19051	32.978455	-117.25003	11.780522

References & Resources

Kevin Crow and William Gould, StataCorp, *How do I graph data onto a map with spmap?*

<http://www.stata.com/support/faqs/graphics/spmap-and-maps/>

Friedrich Huebler, International Education Statistics, *Guide to creating maps with Stata*

<http://huebler.blogspot.com/>

Stata program authors:

`shp2dta` : Kevin Crow

`spmap` : Maurizio Pisati

`geocode` : Adam Ozimek

`vincenty` : Austin Nichols