**Stata Exercise:**

In this example we are going to work through an example of constructing distance measures that hopefully can be useful to you to apply in other cases where you are interested in measuring spillovers that involve a distance component.

There are two datasets:

**gps\_clinics\_mw.dta**, is a dataset of clinics. *Note*: the gps coordinates have been adjusted so they do not actually identify the real location of the data. This adjustment was done to ensure confidentiality of the data source.

**DHS\_GISdata\_mw,** is DHS data from Malawi including the clusterids and the offset gps coordinates.

1. **Setting up the data**
2. Create a unique clinic identifier.
3. Using **gps\_clinics\_mw.dta**, create a new dataset using the **reshape** command that has each clinic longitude and latitude coordinates as variables.
4. Using a **many to one merge**, merge the newly created dataset with the **DHS\_GISdata\_mw.** This combined dataset will be used for the remaining exercises.
5. **Constructing distance measures.**

We want to compute the following distance measures: i) the minimum distance between a cluster and any clinic; ii) the number of clinics within defined radii of a cluster; and iii) if there exists any clinic within a defined radius of a cluster. There are multiple ways to approach this, if you feel comfortable trying your own approach to achieve this – do it! Otherwise try the following approach:

1. Generate a new variable for all 517 clinics that measures the distance between the clinic and the dhs cluster. To do this use the **sphdist** command.
2. Generate a binary indicator variable for all 517 clinics measuring whether the clinic is within 1 km from the dhs cluster.
3. Repeat b for a radius of 2km; 3km; 4km and 5km. Try use a **looping** command.
4. Generate a new variable measuring how many clinics exist within a 1 km radius of the cluster.
5. Repeat d for a radius of 2km; 3km; 4km and 5km. Try use a looping command.
6. Generate a binary indicator for any clinic existing within a 1 km radius of the cluster.
7. Repeat f for a radius of 2km; 3km; 4km and 5km. Try use a looping command.
8. Create a new variable measuring the minimum distance between any clinic and the dhs cluster.
9. Summarize your new variables.
10. **Think of an example where you might be interested in using these types of measures.**
11. **Think of an example where you could use a similar technique for a non-health related project.**