## Day 5: Spatial Analysis

SAR models extend linear regression by allowing outcomes in one area to be affected by

1. outcomes in nearby areas (spatial lags of the outcome variable)

2. covariates from nearby areas (spatial lags of covariates)

3. errors from nearby areas (spatially autoregressive errors)

Download and unzip the shapefile:

<https://catalog.data.gov/dataset/tiger-line-shapefile-2016-nation-u-s-current-county-and-equivalent-national-shapefile>

---------------------------------

cd “Your Folder Path”

-----*Specify the folder where your shape file is*

spshape2dta tl\_2016\_us\_county

-----STATA will create two .dta files from the shape file.

use tl\_2016\_us\_county

----We use the file created by STATA

generate long fips = real(STATEFP + COUNTYFP)

----We generate a fips to identify each geographic unit, postcode .etc

spset fips, modify replace

spset, modify coordsys(latlong, miles)

-----*similar to xtset before panal data regression, we are using the shape file to tell STATA this is a spatial data, note we do it twice, first we link it with the shape.dta file, then we tell STATA what coordinate system and distance measurement we are using.*

save, replace

-----save what we have and prepare to merge it into our data.

---------------------------------------

webuse texas\_ue.dta,clear

merge 1:1 fips using tl\_2016\_us\_county, keep(match)

----*we had the fips as the identifier for each geographic unit, so we use it to merge our dataset with the shape file*

spmatrix create contiguity W

spmatrix create idistance M

-----*There are multiple ways to create spatial weighting matrices, here are two options: W based on nearest neighbors, and M based on inverse of the distance.*

spmatrix dir

----*Then we list the matrices we have created*

---------------------------

Before we do spatial regression, we need a moran test to see if there exists spatial effect.

reg unemployment college

estat moran, errorlag(W)

spregress unemployment college, gs2sls dvarlag(W) ivarlag(W:college) errorlag(M)

estat impact

----*Spatial regression, the coefficients consists of both direct and indirect impact, therefore we want to use estat impact to separate the neighboring effect*