```
cd "/Users/daniellinnen/Box Sync/Spring 2016/B209 TICR III/Final Project"

capture log close

log using "finalproject_log_LinnenD", replace

set more off

version 14

/*	Final Project
	CMS-reported hospital mortality outcomes and California's population
characteristics
	B209
	Daniel Linnen
*/

clear
use "/Users/daniellinnen/Box Sync/Spring 2016/B209 TICR III/Final Project/
LinnenD_final_dataset2.dta"
summarize
describe


//begin renaming variables

rename HospitalName hospital
label var hospital "hospital"

rename DeathsComparedtoNational mortality
label var mortality "Mortality outcome"

rename PopulationPerSquareMile density

rename Medianageyears age

//Destring region values
encode Region, generate(region2)

//Remove measures "too few to report" and "not available"
drop if mortality == "Number of Cases Too Small"
drop if mortality == "Not Available"


// generate new binary variable (worse = 1, not worse = 0)

gen worsemortality = 0
```

```
label var worsemortality "Mortality Below National Rate"
replace worsemortality = 1 if mortality == "Worse than the National Rate"
codebook worsemortality

//turn density into categorical value
sum density
egen densitycat = cut(density), at(0, 1000, 5000, 10000,20000,30000,40000,51000)
label var densitycat "Population Density"
tab region2 densitycat, row chi2

// divide continous pop density by 10,000 for better model interpretation
sum density
gen density10 = density/10000


// Graph density across regions
twoway (bar densitycat region2, sort lcolor(black)), ///
        xlabel(1(1)8, labels angle(vertical) ///
                labgap(minuscule) valuelabel ticks) ///
                        scheme(s2color)

// Descriptive statistics
describe, short
summarize
tab region2 worsemortality, row exp chi2


//Model building
correlate worsemortality region2 density10 age FMR
stepwise, pe(.05) lockterm1 : logistic worsemortality (density10) (age) (FMR)



//Model diagnostic


//test for specification error
logit worsemortality density age, nolog
linktest, nolog


// Linktest is statistically significant suggesting that we are missing the ineraction term
//Add the interaction and run again
gen densage=density*age
logit worsemortality density10 age densage, nolog
linktest, nolog
```

```
// test for linearity
lowess worsemortality age, bwidth(0.5) logit
lowess worsemortality density, bwidth(0.5) logit

// Influential points
logit worsemortality density age densage
predict p
predict stdres, rstand
scatter stdres p, ylab(-4(2) 16) yline(0)

// Goodness of fit
lfit, group(10) table

//final model
logistic worsemortality density10 age


//tables
gen urban_cat = density >1000
tab1 mortality, subpop(MeasureName)
by urban_cat, sort : tab1 mortality, subpop(MeasureName)
mean age, over(urban_cat)
mean FMR, over(urban_cat)
mean density, over(urban_cat)
by urban_cat, sort : table mortality MeasureName
by urban_cat, sort : table hospital, subpop(urban_cat)
by urban_cat, sort : table County, subpop(urban_cat)

save finalproject, replace

log close
```