```
use "H:\oldTS1desktop\ECNS 561\Data Sets\KS_Crime_Data.dta", clear

/*KS Crime Example Problem.
In 1986, the electorate in KS voted to end prohibition on "open saloons"...effectively allowing
counties to vote on whether or not they would allow establishments to sell alcohol by the drink for
on-premises consumption.  The law went into effect July 1, 1987 and, immediately, 36 counties
legalized by-the-drink sales of alcohol.

We are interested in whether or not the crime rate was affected immediately after the law passage.
Particularly, suppose we are interested in whether or not Ci = Crime_1986i - Crime_1988i = 0 where
Crime_1986i is the crime rate in county i in 1986 and Crime_1988i is the crime rate in county i in
1988 for i = 1,...,36 (i.e., the 36 counties that initially legalized by-the-drink-sales.  Assume
the Ci ~ Normal(mu, sigma_squared)

(i) Using data on the 36 counties, construct a 95% CI for mu.*/
*First we need to identify these 36 counties in the data with a variable called "initial changer"
gen law_change_1987 = 0
replace law_change_1987 = 1 if wet > 0 & year == 1987

egen initial_changer = sum(law_change_1987), by(county_fips)

drop law_change_1987

/*Next, we need to calculate the mean crime rate (let's go with violent crime) in 1986 and 1988.
Let's make this easier to look at by just keeping the county-year combos of interest*/
keep if (initial_changer == 1) & (year == 1986 | year == 1988)

*Calculating the difference in crime for each county
gen crime_diff = .

by county_fips: replace crime_diff = violent_rate[_n] - violent_rate[_n+1]


*Now we can calculate the mean and s.d. of the crime difference across these 36 counties
egen mean_crime_diff = mean(crime_diff) if crime_diff != .
egen sd_crime_diff = sd(crime_diff) if crime_diff != .

*With the standard deviation, we can easily calculate the standard error of the sample mean
gen se_crime_diff = sd_crime_diff/sqrt(36)

/*From a table with critical values from a t distribution, we know that the 95th percentile
in the t distribution with df = 35 for a two-tailed test is about 2.03.*/
gen lower_95CI = mean_crime_diff - 2.03*se_crime_diff
gen upper_95CI = mean_crime_diff + 2.03*se_crime_diff



*Let's also look at a 90% CI
gen lower_90CI = mean_crime_diff - 1.69*se_crime_diff
```

```
gen upper_90CI = mean_crime_diff + 1.69*se_crime_diff


/*Q. based on these two confidence intervals, what inference can we make regarding the
statistical significance of our variable mean_crime_diff???
```