

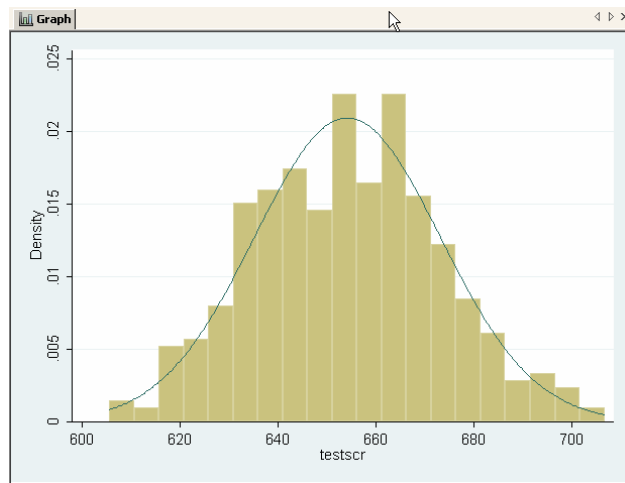
## Homework

The textbook example use datasets of different types: cross-section, time-series, and panel data.

- caschool.dta data. 420 cross-sectional observations on California School district
- macro\_2e.dta 193 time series observations on macroeconomic variables.
- fatality.dta 336 observations.

Import each into Stata and get the summary statistics. What is the average test score? What is the average student teacher ratio for classes with test scores above 650? Below 650? Plot a histogram of **testscr** and add a normal density plot on top of it (**Graphics>Histogram**).

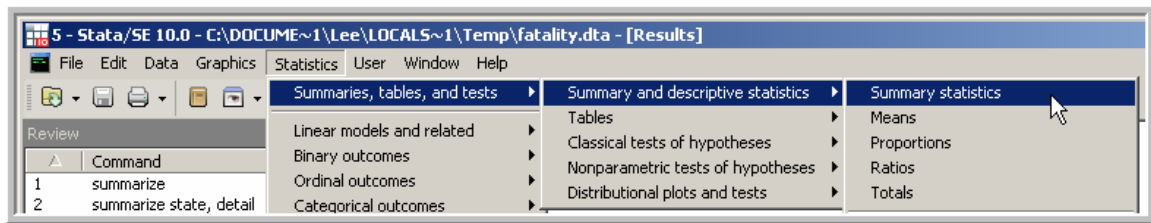
It will look like this:



Declare the macro\_2e data to be time series using **Statistics>Time series>setup and utilities>Declare dataset to be time-series** dialog. What years are covered? Generate the natural log of **punew** (price level). Use the difference operator **D.** to take the differences of this variable and call it inflation. Plot inflation against time using time series plot. In what time period did inflation reach a peak? (Hint: **sort** and **list** or use an **if** statement).

Declare the fatality data to be panel data using **Statistics>Longitudinal/Panel data>setup and utilities>Declare dataset to be panel** dialog. What years are covered? How many states are in the sample? Is the panel balanced? What is the average beer tax for each year? Has it been increasing or decreasing?

To get this use



and use the conditional **by** statement available from the **by/if/in** tab.