In earlier labs, we used data that were already *Stata* friendly. We recognize those files since their filenames end in .dta. What if we want to use our own data? In this lab, we will enter medical data, which were collected by Christina Palmer in a recent study of use of Captopril, a medication for patients suffering from high blood pressure.

**A Brief Description of the Captopril Experiment**

In order to test the effectiveness of the drug Captopril as a blood pressure treatment, fifteen patients were given equal doses of the drug. For each patient, the first variable represents the systolic blood pressure before Captopril was administered, while the second variable is the systolic blood pressure after the patient took Captopril.

```
var1 210 165 187 160 167 176 185 206 173 146 174 201 198 148 154
var2 201 165 166 157 147 145 168 180 147 136 151 168 179 129 131
```

To begin, we type
```
. edit
```

in the Command window and press enter. We enter and hit return after each value of var1 and var2. We close the *Stata* Editor window by clicking the red button in the upper left corner.

To replace the title var1 with “before” and var2 with “after”, we type
```
. rename var1 before
. rename var2 after
```

Next we want to summarize the two variables.
```
. summarize before after, detail
```
Question 1: Compare the average systolic blood pressure for patients before and after they took Captopril.

. graph before
. graph after
. graph before after, box

Question 2: Describe what you see when you plot these data.

Since these data are paired, with two measures for each patient, we want to look at the relationship between the measures. A scatter plot might be helpful.

. graph after before

an then

. graph after before

Question 3: Which variable should we treat as the independent variable? With some paired data, it does not matter which variable we place on the horizontal or the x-axis. However, in this case, it seems more reasonable to place “before” on the horizontal as the explanatory variable. It seems reasonable that “after” is our response or dependent variable and should be placed on the vertical or the y-axis.

Since we are interested in the differences between the variables, we want to create or generate a new variable for those differences.

. generate diff = after - before
We can plot these differences using a histogram or box plot, but a stem-leaf plot would also be useful.

. stem diff

Question 4: If Captopril is ineffective, then there would be no real difference in the “before” and “after” values, then where would we expect the center of the plot of differences to be? Does that seem to be the case? Is there evidence from the summary statistics and plots, that using Captopril decreases the systolic blood pressure of the patients in our study?

Assignment for Inputting Our Own Data Lab

Your TA will tell you when the assignment is due. Here is a brief description of data collected by Matash Esfandiari. A sample of twenty-seven children ages 8-9 years old who suffer from a condition known as transient tyrosinemia (TNT) were paired with a control group of twenty-seven children who did not suffer from TNT. The Illinois Test of Psycholingual Ability was administered to each child. We will focus on the auditory reception score, one of ten test variables. The score for each pair of children is recorded in the following table.

| Con Group | 40 35 30 22 21 39 39 22 44 34 30 26 44 36 30 |
| Exp Group | 26 31 28 19 31 35 37 41 35 29 18 18 38 31 26 |

| Con Group | 18 27 30 33 26 31 29 34 36 42 32 38 |
| Exp Group | 33 27 22 37 24 11 24 28 17 31 41 23 |

Do the data suggest that there is a difference in the auditory reception scores for children who suffer from TNT and those in the control group? Please justify your answer by using techniques from this lab or previous labs.