Open up the data set that goes with this assignment (‘Assign8.sav’). It contains data from variable Y, a cardinal variable that uses a score of -1 to represent ‘no response’.

Go to the Edit>>Options menu and click on the Viewer tab. If the ‘Display Commands in the Log’ option isn’t already selected then select it, this will cause SPSS to print the commands you give it along with the output that results from those commands. It probably already is selected, as that is the default setting. In this lesson we will take a look at those commands. But first, let me show you how to ‘select cases’. Close the Options window.
We want to look at the mean and standard deviation of the variable, but the presence of the ‘no response’ scores of -1 will foul up the computations (we are only interested in the people who did give a response, and the use of -1 as a label for no response turns the variable into a nominal variable, making the mean meaningless).

Go to the Data>>Select Cases menu. This is a very useful place to know about in SPSS. We want to tell SPSS to filter out all scores of -1 and keep those that are legitimate. Select the ‘if condition is satisfied’ option, and click the ‘If’ button, then in the upper right hand box of the ‘Select Cases: If’ window type in Y > -1 (so we select only those scores that have real values and ignore those that equal -1), then click ‘Continue’.

Now we have an important decision, do we want SPSS to simply ignore ('filter out') the scores of -1 when doing the analysis, or do we want SPSS to create a new data set that only contains the cases that we want, or do we want SPSS to actually remove those scores from the data set? I don’t want to change the data set or create a new one, I just want to be able to pick which scores to analyze, so select ‘Filter out unselected cases’, then click ‘OK’.

Go back and look at the data. Notice how SPSS has indicated which scores will not be analyzed by putting a line through the row number. Also notice that SPSS has created a new variable that has a ‘1’ for all of the scores to analyze and a ‘0’ for those to not analyze. You can ignore that variable.
Now, close the output window (don’t bother to save it) so that the next time you ask SPSS to do something it will bring up a nice clean window, this isn’t necessary but it will make what happens next clearer.

Go to Analyze>>Descriptive Statistics>>Frequencies, click on the Statistics button and indicate you want the mean and standard deviation. If you haven’t already, move variable Y over to the ‘Variable(s)’ box, then click OK.

Because you earlier indicated that you wanted the log to be displayed on the output, the following text will appear above the actual output of the analysis you just requested.

FREQUENCIES
   VARIABLES=Y
   /STATISTICS=STDDEV MEAN
   /ORDER=
ANALYSIS .
What is that? Before the Macintosh revolutionized the computer-user interface with its use of the mouse, menus, and buttons, people using SPSS would have to type in the lines given above to get the program to find the mean, median, and mode of variable Y. These commands are part of what SPSS calls ‘syntax’, which is the computer language you use to need to know to tell SPSS what to do. Those commands are still in SPSS, but now selecting from menus and pressing buttons on the screen creates those commands for you and sends them on into the inner workings of SPSS, and unless you ask for the data log, or use the syntax yourself, you never see this code or even need to know that it exists.

For the vast majority of the work we will ask SPSS to perform we can ignore the existence of SPSS syntax code, we will just select from menus and click on buttons and view the results. There are some good reasons, however, to learn how to work with syntax, so I wanted to introduce it now. We will just dabble in it this semester, but that should be enough to give you an idea of what it adds to SPSS and will serve you well if you need to do syntax beyond the confines of this course.

OK, enough about syntax. If you successfully filtered out the ‘no responses’ then your Frequencies output should indicate that you had 30 valid scores. The data set has 33 scores but you filtered out the ‘no response’ scores before you had it compute the mean and standard deviation.

24. Mean=_____________

25. Standard deviation = _____________
Open Prob1.sav in SPSS. Now go through the following steps:

a. Variable Y is reaction score times in msecs, let’s label them as such, but I want to do it a slightly different way than going to the ‘Variable View’.

So instead, go to Data>>Define Variable Properties, move Variable Y over to Variables to Scan and click Continue. When the Define Variables Properties window appears type ‘Reaction Times (msecs)’ into the label box and click OK (This step will be explained later explain why I’m underlining this part later).
b. Now, let’s take a look at the reaction times. Go to **Graphs>>Legacy Dialog>>Histogram**, move Reaction Times into the Variables box and **click OK**.
c. Look at the reactions times, reaction time data are typically skewed and so are these. With reaction times, taking the inverse of them (1 divided by the reaction time) often results in scores that are more normally distributed. Let’s check it out.

Go to **Transform>>Compute Variable** and create a new variable called ‘InverseY’, in the Numeric Expression box put in the formula ‘1/Y’, and **click OK**.
Check your data to make sure it created the variable. You will see mainly .01’s and .02’s simply because only two places to the right of the decimal point are being shown, if you asked SPSS to show more to the right of the decimal you would see that actually the inverse scores differ quite a lot.
a. Now let’s see what those inverse scores look like. Go to **Graphs >> Legacy Dialog >> Histogram**, remove Reaction Times and put InverseY into the Variables box and click OK.

The InverseY data look at least slightly less skewed and more normally distributed than the original Y data. Interesting, huh? (maybe I’m easily amused).
Ok, now. Close the data set without saving it. Open it up again and do exactly the same thing again (‘a’ through ‘d’ above) but each time when you get to a click OK, instead of clicking Ok click ‘Paste’. The first time you do it a ‘syntax’ window will appear with the relabeling instructions.

Don’t close it, just move it to the side and go on to part ‘b’, when you press ‘Paste’ instead of OK note that the histogram instructions are added to the end of the relabeling instructions. Go on to part ‘c’ and ‘Paste’ at the right time.

You might notice that SPSS isn’t doing anything you have asked it, it is simply putting the instructions into the syntax window instead. Go on to part d...whoops, there is a problem, and you can’t indicate you want a histogram of InverseY because it doesn’t exist yet. No problem, look at the syntax so far:

```
*Define Variable Properties.
*Y.
VARIABLE LABELS Y 'Reaction Times (msec)'.
EXECUTE.
GRAPH /HISTOGRAM=Y.
COMPUTE InverseY = 1/Y.
EXECUTE.
```

The part that says to do a histogram is:

```
GRAPH
/HISTOGRAM=Y.
```

Copy and paste that to the end of the syntax, and edit /HISTOGRAM=Y to /HISTOGRAM=InverseY, now your syntax should look like this:

```
* Define Variable Properties.
*Y.
VARIABLE LABELS Y 'Reaction Times (msecs)'.
EXECUTE.
GRAPH /HISTOGRAM = Y.
COMPUTE InverseY = 1/Y.
EXECUTE.
GRAPH /HISTOGRAM=InverseY.
```
Close the output window so that you will get a fresh one when you do the following (making it easier to see what happens).

Now, finally, to get SPSS to do these instructions go to the Run menu in the syntax window and select ‘All’. SPSS will now do all of the instructions in one fell swoop. You should have exactly the same output you did when you did it all the first time.

So, what’s the big deal about syntax? Open Prob2 and then close Prob1 (do it in that order or the syntax window will be erased...SPSS closes down completely when you shut all data windows). Prob2 is also a set of reaction times labeled as Variable Y. Assuming that you still have your syntax window with its code in it, go to the Run menu of the syntax window and select ‘All’ again. Presto, it just did everything the same way without you having to do any real work! Once again look at the two histograms, the reaction times are rather skewed and the inverse reaction times are rather more normally distributed.
So here is the biggish deal about syntax.

1) You might have several, perhaps many, steps you want to perform on the data, and then you are going to want to do the exact same steps on further data (perhaps collected in the future or perhaps existing now in different data sets).

Paste all the steps into the syntax window, and you can then save the syntax, it gets stored as a ‘.sps’ file. Any time you want to analyze data the same way, load up the data, then open up the syntax file, and run it. If the variables are called something different than just change their name in the syntax code (e.g. replace ‘Y’ with ‘X’).

2) If someone else has written a syntax file that does what you want to do, get a copy of it and open it up in SPSS, and presto you can do it too. There are syntax files available on the web, you can just cut and paste them into the syntax box.

3) By having SPSS include the syntax in the log on the output (we covered that last assignment), when you come back a year later and try to remember what in the heck you did, there it is.

For all of that....I’m not a big fan of syntax, but then I’m not doing the sort of research that requires a long string of complicated steps that has to be recreated each time. Still, it does open up what you can do with SPSS and I wanted you to know about it.

Oh yes, something to input into the homework for credit. So, look at the histogram from the InverseY scores from Prob2. The first number from the left on the horizontal axis is 0.008

37. The second number from the left on the axis is .010
38. The third number from the left on the axis is .012
39. The standard deviation of the distribution (shown to the right of the histogram) is .003