SocSci 200: Social Science Perspectives on the Family
Spring 2000
Introduction to SPSS Lab
May 11/12, 2000

First task: Start up

1. Click the icon on the very lower left-hand side of your screen, labeled "Start" (with the Microsoft logo).

2. This will throw a menu on the screen. If the previous user of your computer did not logoff, you will want to shut down the connection and start again. In this case, move the cursor to "shut down". This will throw up another set of options. Select "close all programs/start as different user," and click.

3. The screen will now ask you for a password to the login name "public." There is no password for this account, so just leave the password field empty and hit return.

4. The computer will take a few seconds to log you in. Once the hourglass disappears from the screen, you can proceed.

5. Click "Start" again, locate "Run" on the menu of options and click. Windows will ask you to type in the name of the folder or file you want to run.

Type the following: o:classdata ssci200. This tells your computer where to go to get the data you'll analyze with SPSS.

6. Locate the "SPSS for Windows 9.0" icon on the desktop and click.

7. Now the computer is entering SPSS, otherwise known as the Statistical Package for the Social Sciences.

Second Task: Loading the Data Set:

1. Your screen might be titled "SPSS for Windows Viewer" and contain a warning. Don't be alarmed; it simply means that SPSS is looking for data. Ignore this warning and move to the 'SPSS Data Editor' window by clicking on the "Untitled-SPSS for Windows" button at the bottom of the screen. The window that pops up should be empty and look like a spreadsheet. What we need to do is retrieve some data and load it into this window.

2. In the upper-left of the Data Editor screen, click on 'File'.

3. The 'File" menu should pop down.

4. Click on 'Open'.

5. The 'Open File' window should pop up.

6. The prompt will be "Look in:" Find the C: drive and click.

7. Find the folder labeled 'Temp' and click.

8. Find the folder labeled 'ssci200' and click.

9. Now locate a file titled 'GSS93 subset.sav' and click

10. Click the Open button.

11. You should see all kinds of stuff on the screen!

Third Task: Looking at the data.

1. The stuff on your screen is data from the 1993 release of the General Social Survey (GSS). The GSS is an annual survey of the behavior and attitudes of English-speaking persons 18 years or older, living in non-institutional arrangements within the U.S. The far left column, "id," contains randomly assigned case (people) numbers. There are 1500 people included in this dataset. Each case represents a person who was surveyed in the GSS. All of the other columns contain, for each case, the values for social scientific variables.

   For example, find column 3 "marital." The number in each cell in column 3 is a code that corresponds to the marital status for that person (case). To find out the codes, left double-click on the column. You'll get a window titled "Define variable." Then click on "Label." You'll get a full description (not just a name) of the variable. You'll also discover that '3' means 'divorced,' '5' means 'never married,' etc. Clicking on 'missing values' enables you to control which codes are defined as missing (you won't want to do this now, however). When you have finished learning about the variable 'marital,' click 'continue,' then 'OK.'

   Use the arrows on the outside of the window to play with the data set: scroll up and down. Choose a variable name (column) that looks interesting to you and left double-click on it to discover what the variable is and how it's coded.

In a short paragraph , describe what you've found about this variable.

Fourth Task: Getting Measures of Central Tendency

1. At the top of the Data Editor screen is a toolbar with words like Data, Transform, Analyze, Graphs, etc. These are your menus: each menu provides you with a set of statistical or data manipulation tools. Click to get an idea of what each menu has to offer.

2. Click on the "Analyze" menu at the top of the screen. Then click 'Descriptive Statistics.'

3. Click 'Frequencies.'

4. Find the variable 'agewed' on what is akin to a shelf of variables and highlight it. You can use your mouse to move the 'Frequencies' window out of the way; then left double-click on the column to find out more about this variable.

5. Go to the 'Frequencies' window and click on the arrow between the shelf of variables and the 'Variables' box.

6. Click on the 'Statistics' button at the bottom of the 'Frequencies' window.

7. The 'Frequencies: Statistics" window should pop up. This window allows you to calculate measures of central tendency for the variables you highlight.

Measures of central tendency: A measure of central tendency is a statistic that, for a given set of data, represents the typical or "average" value for a variable. We will be dealing with three different measures of central tendency: the mode, the median, and the mean.

The **mode** is defined as the most frequently occurring value for a variable in a set of data.

The **median** is defined as the value that divides a set of scores in half: 50% of the cases have values that lie below the median, and 50% of cases have values that lie above it.

The **mean** is a mathematical quantity that can be thought of as a weighted average. It is computed by adding up the values of all the cases and dividing that sum by the total number of cases.

Mean = Sum of Values for Variable/ Number of Cases

Example: You want some sense of the "average" score on a midterm exam for a class of 21 students. The variable is "score on the midterm exam." Below are the scores; they are organized from the lowest to highest value because that is how one finds the median.

23 25 37 37 41 44 56 59 63 72 73 75 75 75 79 82 86 88 91 95 95

The **mode is 75** because it occurs three times, more than any other value in the series.

The **median is 73**. This score is in 11th place in the series, which is the midpoint for this distribution of 21 scores.

The **mean is 1371/21, or 65.29**.

Note that each measure of central tendency gives a different sense of the "average." This is especially true of the mean, which is quite a bit lower than either the mode or the median.

For any distribution of values for a variable, the mode, median, and mean can turn out to be either very close together or quite different. When the **mode, median and mean are equal or nearly equal to each other**, this means that the **overall distribution is symmetrical about the "average."** When the **mode, median and mean are different**, as in the above example, this means that **the distribution is skewed.** In skewed distributions, values for the variable are either concentrated toward the high end (with only a few low values), or concentrated toward the low end (with only a few high values). In the midterm example, the mode, median, and mean indicate that we have a skewed distribution. Looking back at the original data, we can see that, in fact, most of the midterm scores range between 72 and 95 (the high end), with relatively few low scores.

Let's return to the computer and put this knowledge to work:

8. Run the appropriate measures of central tendency for the variable 'agewed.' Do this by finding the box "Central Tendency" and clicking the boxes left of the mode, median, and mean.

9. Click the "Continue" button.

10. Click on the 'OK' button.

11. Now the computer will make number-crunching sounds...

12. In a new window, you should see your measures of central tendency for the variable 'agewed'; SPSS also automatically prints a frequency distribution for the variable. To move about the output in SPSS, you can either use the mouse to click on different tables/figures shown in the tree diagram on the left-hand side of your screen, or you can scroll up/down using the right sidebar.

You can also type directly onto the output by clicking on Insert (toolbar) and then New Text. Use this technique to answer the following questions:

Give one-sentence interpretations of the mode, median, and mean for "agewed."

Is this distribution symmetrical or skewed? How can you tell?

13. **Print a copy of your output/answer by clicking the printer icon.** Return to the tree diagram of your SPSS output (left-hand side of your screen). Click on Frequencies; this will select all components of your output for printing. Then click the printer icon on the toolbar.

14. The Print window will pop up. Click 'OK.'

15. Put away your output by double-clicking the button in the upper left-hand corners. Answer 'no' when SPSS asks if you want to save your output.

<u>Fifth Task: Saying Goodbye.</u>

1. Call up the Data Editor window if it is not already on the screen by clicking on the "GSS93 subset" button at the bottom of your desktop.

2. Click on 'File' at the top of the screen.

3. Click on "Exit."

4. You're done! Congratulations!

Save this handout; it will prove useful in upcoming lab sessions. And don't forget to turn in your question answers and printout for class participation points!