SPSS Lesson 1
Variable Definition

Note there are two different "views" you can use in SPSS – "data view" and "variable view." You can see the toggle switch for these at the bottom left hand corner when you open up SPSS.

You will be entering your data in "data view," but the place to begin is with "variable view" so that you can define your variables.

There are two main reasons you want to fully define your variables:

- To keep an electronic archive of your data and all its attributes (e.g., complete survey items; coding scheme)
- To set up your data so that when you start to run analyses, the results and any tables and graphs you produce will be attractive, complete, and understandable to the reader.

The following elements are important:

1. NAME refers to "variable name."
   - In row 1, you would enter the name of your first variable.
   - This name must be 8 characters long or fewer, and the first character must be a letter.

2. TYPE refers to the type of variable. When you click on TYPE, a dialogue box will open in which you need to specify the type of variable you have. The main two we will deal with are "string" and "numeric" variables.
   - STRING variables include letters, and can be used for names or perhaps brief open-ended responses.
   - NUMERIC variables are numbers. In either case, you'll need to specify your variable width. You can do that in the dialogue box, or in the subsequent WIDTH and DECIMAL columns.

3. Since variable names often tend to be very cryptic (because they must be 8 characters or less), LABEL allows you to specify a longer variable name. This longer label will appear on any charts or graphs you produce.

4. VALUES allows you to connect the values (numbered codes) of your coding scheme to the original categories. For example, it is here where you tell SPSS that for the variable SEX, males were coded with a "1" and women with a "2." When you click the square, and then click on the button with the three periods, a dialogue box will open. In the square titled "value" place the first code, e.g., "1". In the square entitled "value label," place the category label, e.g., "male." Then click "add" and see the number and category appear in the lower box (e.g., 1 =

"male"). Continue doing this until you have entered all the different codes you used for that variable.

5. MISSING refers to a missing data code. This is a "special" number that SPSS will treat as a unique code to identify places where you do not have data. The upshot of doing this is that SPSS will avoid including it as a "real" number when statistics are computed. Note that when you click "missing," a dialogue box appears. Note that more than one missing data code is permitted for each variable, just in case you want to code the different reasons values might be missing, e.g., 8 = not applicable; 9 = should have answered but didn't.

6. COLUMNS refers to how many columns wide you would like the variable to be presented in the "data view." Normally this would be at least 8 so that the variable name could appear easily. The subsequent column, ALIGN, allows you to change the presentation so that scores for the variable are left-justified, right-justified, or centered.

7. MEASURE refers, for numeric variables, to the level of measurement of your variable. It distinguishes three: nominal, ordinal and "scale" variable.
   - "Nominal" variables are also known as "categorical variables;" the different codes here refer to different categories that [e.g., SEX is a categorical variables because the two different groups are simly different categories that bear no mathematical relation to each other).
   - "Ordinal" refers to variables where the numerical codes reflect an ordering of some sort, but where the distance between the categories can vary. For example, "academic rank" is an ordinal variable – the professorial ranks are ordered (1) assistant professor; (2) associate professor; and (3) full professor – but the distances between the codes are not necessarily equal.
   - "Scale" variables include interval and ratio levels of measurement, where any numeric codes have meaning in terms of number relations that go beyond "mere" category and order; if it makes sense to compute an "average" (or mean) score for the variable, then you are probably dealing with a "scale" variable.