Introduction to Quantitative Sociology
SPSS Assignment 3

Due: Tuesday, November 19th in class (Note: This is a change from the syllabus)

In this assignment you will be learning how to run and interpret bivariate SPSS analyses. I am assigning you relationships you will have to analyze, based on which dataset you are using. This will help you to prepare for the take home final, in which you will choose your own relationships to test out.

Student Dataset:
If you are using the student dataset, you should analyze relationships between the following three sets of variables:
1. Minority Status and Smoking
2. Sex and Diet
3. Diet and Health

Crime or Health Subset:
If you are using the Crime or health dataset, you should analyze the relationships between the following three sets of variables:
1. Minority Status and Marstat
2. Sex and Family Income
3. Education and Family Income

No matter which dataset you are using, you will need to list a research hypothesis for each pair of variables. In other words, you need to specify what relationship you expect to find between these two variables. You should do this BEFORE you run your data analysis. They do not have to be correct in the end. You will have three research hypotheses in your paper.

No matter which dataset you are using, you will need to draw a path diagram which depicts each research hypothesis. You should do this before you run your data analysis. You will have three path diagrams in your paper.

No matter which dataset you are using, you will need to compute a new variable called MINORITY. It should be equal to 0 if the respondent is white and not Hispanic, 1 if the person is not white and/or hispanic. Instructions for doing this are included below. (Note: If you saved your data when you did this for the 2nd assignment, then you will not have to redo it.)

You should decide what kind of bivariate data analysis you will need to conduct to test each of your hypotheses, taking into account the kinds of variables you will be using. Instructions for conducting each of the three kinds of statistical analysis we have learned are included below.

You should write up the results of your analyses in complete sentences and whether each of your hypotheses are supported. In this write-up you should make it clear what kind of statistical analysis you have conducted, and report the appropriate statistics in your sentences.

Your paper will likely be 1-2 pages. You will be graded on your appropriate wording of hypotheses, whether your path diagrams reflect your hypotheses, whether you used appropriate bivariate techniques, whether your write-up accurately reflects the results of your analysis, and whether you come to the correct conclusions regarding your hypotheses. Your paper should be typed, double-spaced, and in a standard 11-12 point font. Grammar and spelling will be 10% of your grade.
THINGS YOU WILL/MAY NEED TO BE ABLE TO DO IN SPSS

Opening your data, running frequencies, and assigning missing values: These were all described on previous assignment sheets. Go back to those sheets for directions for these things (or find someone who still has a copy).

Recoding/Computing New Variables: You will need to compute a variable for minority status no matter which data you are using. You may also want to compute other new variables that recode some of the original variables in new ways.

To compute a new variable, first make sure you are in the data view window (not output or variable view). Then click on Transform, then Recode, then Into Different Variables. (Note: Although you can recode an existing variable, we generally avoid doing that so that we don’t screw up our original variables.) You will now be in the Recode Into Different Variables dialogue box.

If you saved your data after the last assignment, you should already have the minority variable in your data. If not, the instructions to construct this variable are available in SPSS Assignment 2.

Bivariate Analytical Techniques

Running Crosstabs: If you need to run crosstabs, click on Analyze, then Descriptive Statistics, then Crosstabs. Find your dependent variable in the list of variables to the left, click on it once, and then click on the arrow next to the Row(s) box (the top empty box). You should see your dependent variable in the Row(s) box. Next click on each of the demographic or independent variables you need crosstabs for, and click on the arrow next to the Column(s) box. You should see each of them get added to the Column(s) box. Now click on the Cells button. Under Percentages, click on Column and Total. Then click on the Continue Button. Now click on the OK button. You will have a series of crosstab outputs in your Output file (1 for each independent variable you chose).

Running Difference of Means: If you need to run difference of means tests, click on Analyze, then Compare Means, then Independent Samples T-Test. Find your dependent variable in the list to the left, click on it, and click on the arrow next to the Test Variable box. You should now see your dependent variable in the Test Variable box. Now choose one of the independent variables you need to run a difference of means test with. Click on it and then click on the arrow next to the Grouping Variable box. Now click on the Define Groups button under the Grouping Variable box. Fill in the values for two categories of the independent variable in question. (For example, if you were using Minority Status for any of the datasets, the two values would be 1 - Minority and 0 - Non-minority.) Now click on Continue and then click on the OK button. You will now have Difference of Means results in your output file. The means for each group will be listed in the Group Statistics box.

Running Correlations: Click on Analyze, Correlate, then Bivariate. Double click on your Dependent Variable first, so that it moves into the empty Variables box. Then double click on any of your independent variables that you want to run Correlations with. Once you have all the variables you need in the Variables box. Click on OK.

You will now have a Correlation matrix in your output file. The top number in each cell is the correlation coefficient for the relationship between the two variables.