By using an existing data set, this lab is intended to familiarize you with some of the basic elements of the research process. Using a specific variable as a jumping-off point, you will be asked to formulate a plausible research question, develop suitable explanations, identify the concepts associated with those explanations, and formulate testable hypotheses.

1. Start STATA, download the Cigarette.dta data set, and open a log file

In the command window, type

[a] \texttt{use http://web.ics.purdue.edu/~ewaltenb/POL300/Data/Cigarette.dta}

Notice the command you just typed appears in the Review window and variables are listed in the Variables window.

[b] Open a log file

2. Examine the sales variable

type \texttt{des sales}

! a. What does sales represent? This is the concept the sales variable is representing.

___________________________________________________________________________

___________________________________________________________________________

___________________________________________________________________________

b. Formulate a testable research question that is based on the sales variable. (Note: since you are working with a finished data set, your question must be answerable with the data at hand. That is, you should be able to attempt to answer your question with the other variables present in the Cigarette.dta data set.) ________________________________

___________________________________________________________________________

___________________________________________________________________________

___________________________________________________________________________

___________________________________________________________________________

3. Formulate testable hypotheses

! a. For the question you just formulated, and bearing in mind you must attempt to answer it with the Cigarette.dta data set, come up with 3 testable hypotheses.

9 H1 ________________________________________________________________
b. For each of these hypotheses, identify the dependent and independent variables.

H1 [IV] _________________________ [DV] _________________________________

H2 [IV] _________________________ [DV] _________________________________

H3 [IV] ________________________ [DV] _________________________________

c. Identify the type -- correlative, directional, casual -- of each of your hypotheses.

H1 ___________________________________________

H2 ___________________________________________

H3 ___________________________________________

d. Identify the concept the independent variable in each of your hypotheses is representing.

H1 __________________________________________________________________

H2 __________________________________________________________________

H3 __________________________________________________________________

e. What is the unit of analysis that these data and your hypotheses are examining? _________
4. For fun, test your hypotheses

Something we will learn in a couple of weeks is "levels of measurement." They are important, because the level of measurement at which your data are collected affects what statistical procedures are appropriate in their analysis. Nearly all of the variables in the Cigarette data set are measured at the ratio level. This is highest level of measurement there is, and it permits analysis using a procedure called Pearson's Correlation. The procedure measures the degree to which 2 variables move or vary together. The greater the degree to which they move together, the stronger or closer is their relationship. Pearson correlations vary between -1 and +1. The closer they are to |1|, the stronger the relationship. The closer the correlation is to 0, the weaker the relationship. A negative correlation means that the 2 variables are moving in opposite directions, while a positive correlation indicates that 2 variables are moving in the same direction.

To conduct this procedure, type `pwcorr sales` and then the 3 independent variables you identified above. Were any of your hypotheses correct? _________________________________________
______________________________________________________________________________
______________________________________________________________________________
______________________________________________________________________________

There's a lot more to the interpretation and work with these correlation coefficients, and we'll work get into that after the first exam.

5. To Turn in

The completed lab. Submit it and your log file. Be sure that your name and SS# are on all the material you submit.