

“Lab Practices”

I. Always keep a lab notebook. Keep track of ...

A. Hypotheses to be examined

B. All transformations of your data

C. All Analyses you perform.

II. Name each new variable you create. Use names that are substantively meaningful if possible. (This is less easy to do in large data files where name creation can take a long time.) Be consistent in your naming style. Use all upper or all lower case letters.

III. Label each variable you create. The label gives you (and anyone else who might use your data) the opportunity to further understand what concept a variable measures. This is obviously critical if your variables are simply named with numbers such as “VAR001”, “VAR002”, etc. But it’s also true if your variables have short mnemonic names that may be meaningful only when the data were created but whose meaning is difficult to create later. “PREZAPP” may mean “Presidential Approval Score” and perfectly plausible but it’s not obvious what PREZAPP is if you look at the variable name in isolation.

IV. Code and recode each variable so that it is easiest to interpret in light of your research hypotheses. If you believe, for example, that political conservatism is related to support for the Republican candidate, make sure that the political ideology scale is coded with “extreme conservative” as the highest value and “extreme liberal” as the lowest value AND that vote is coded with

the high value for the “Republican” candidate and the lowest value represents the “Democratic” candidate. In this coding scheme, high values should go with high values and low values with low values.

V. For all categorical data, label each value of each variable you create. Tables will look much better and be more interpretable if you have value labels.

VI. Remember you must assign or reassign value labels when you recode variables. If you recode a variable into the same variable name, make sure that the new values have appropriate labels. (Sometimes this means recoding all values but sometimes only partial reassignment of value labels is necessary.) If you recode an existing variable into a new variable you must assign value labels to all values.

VII. Always examine the frequencies of a variable you have transformed.

A. If “recode”, did the recode produce the result you expected?

B. If “compute”, do the values of the new variable look like they should?

C. Verify that missing data is handled correctly on any recode or in the creation of new variables.

VIII. Whenever you create a new version of the data, save it in a separate file (under a different name). Consider using a version number as a suffix. For example, label the first file as “data_1.sav”, the second file as “data_2.sav”, and so on. Keep records of every change you make that leads from one version of the data to the next.

