Data entry and codebook creation

Purpose

Although you will be working with previously collected data, it is important to understand what data looks like as well as how it is coded and entered into a spreadsheet or dataset for analysis. This can help you identify and avoid problems later when reading data into an analysis software program. For example if you mix letters and numbers in the same cell, the variable will be treated as character not numeric.

There are three pieces to this assignment.

1. Entering raw data into a spreadsheet.
2. Creating a codebook
3. Importing the data into your software program of choice (SPC)

Instructions

- Using the PDF copies of medical records for 5 patients seeking treatment in a hospital emergency room you are going to do data entry and create a codebook.
- You will enter data and create your codebook directly in Google Sheets.
  - Start a new spreadsheet in the **01 Data Entry** folder in our shared Google Drive.
  - Name this file `medrecords_userid` where *userid* is your chico state user id.
  - Name the first three worksheets within this file: `data`, `codebook`, `import`.

Data Entry

1. Select 4 variables recorded on the medical forms
   - one should be a unique identifier, at least one should be a quantitative variable and at least one should be a categorical variable
2. Select a brief name (ideally 8 characters or less) for each variable - write this in the first row
3. Determine what range of values is needed for recording each variable
4. Enter the data for each patient, one patient per row.
5. If data is missing for a particular value, leave the cell blank.
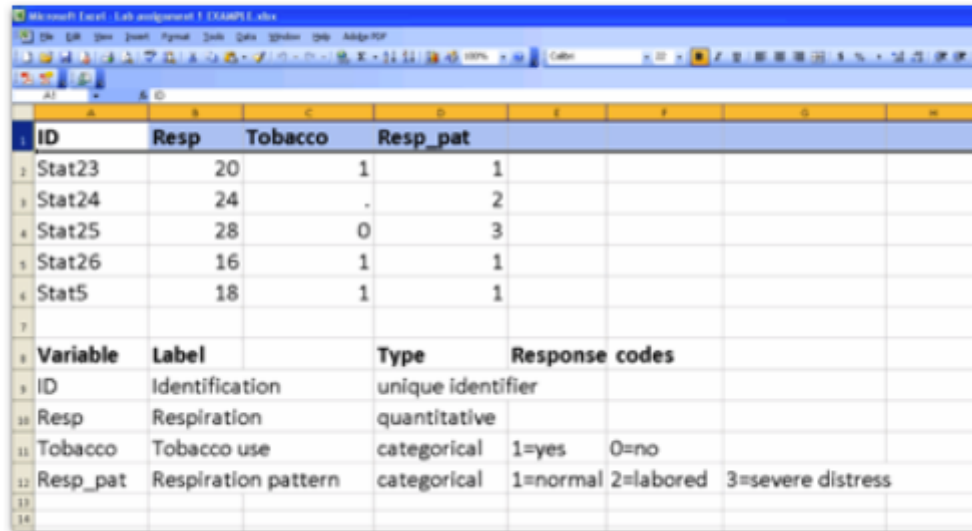
Codebook Creation

In a separate worksheet list the variable names, labels, data types, and response code or ranges in separate columns (4 columns total).

An example of what this should look like is below. The only modification is that your codebook should be on a separate worksheet (not at the bottom like this one shows)

Data Import

1. Using your software program of choice, import this data into the program using point and click GUI methods.
   - Code is fine if you already know how. Point and click is just for the true newbies.
   - Save your file to your hard drive as a Comma Separated Value (`*.csv`).
   - R Studio: https://support.rstudio.com/hc/en-us/articles/218611977-Importing-Data-with-RStudio
   - SPSS: https://libguides.library.kent.edu/SPSS/ImportData
2. Note and record any problems that you noticed and/or had to fix at the bottom of your `codebook` worksheet.

Model:

Figure 1:

- Does your data file look like your spreadsheet?
- Did you have to specify missing values in any specific way?

3. Copy the code that is written for you when you import your data. Include this in the `codebook` worksheet at the bottom.

- This may pop up in another window, (SPSS users have to specify to **paste this syntax**)
- This is a critical part of the assignment.