# Exporting the Data; Keys & Values

Insofar, we have seen our survey data includes some numerical keys which map to a defined text and performed this very mapping on our data set last time. We have also seen a tremendous number of missing values and explored the reasons why these may be present. Now, we will take the data, both with and without mapping, and, both with and without missing values, and add these separate dataframes into an excel workbook as separate sheets, fixing problems along the way.

# The Column-Problem

Noted before were the lengthy variables; the column headers which often would include entire sentences. These long column names are quite valuable, as they display the question or request to the participant. Therefore, the chosen solution to this included MultiIndexing. MultiIndex columns in pandas allows for a sort of "parent" variable for a given set of child variables in the header. Because of our purposes of accessing a variable's contents quickly, each child column in the data was assigned its own parent variable including its type (metadata or question) and number.

# DataFrame Collecting

Now that we have applied more concise variable names to our data, we grab the dataframes we need for excel. The suffix "\_text\_mapping" was applied to the recent dataframe and two others were pulled from this: one containing no missing values and one containing only missing values.

Our input data for Week 2's script had gone through some processing but was never mapped to a text. For this reason, we grab this data, perform all the processing from Week 2 excluding the mapping, and follow up by performing all the same processing done to the data set above.

# Creating the Workbook

We now have the following dataframes:

\* Survey Data w/ Keys

\* Non-Missing Data / Keys

\* Missing Data w/ Keys

\* Survey Data w/ Text

\* Non-Missing Data w/ Text

\* Missing Data w/ Text

We then create a workbook to hold these data sheets for analysis. Cleaned data is now readily available.