PSYCH 167 Jupyter Error Guide

For all errors, scroll *all the way down* to the bottom of the printed error to see the main cause.

- NameError: name 'Table' is not defined; NameError: name 'my_table' is not defined
  - common cause: All cells need to be run in order, and do not skip the top one that imports all the libraries
  - This error means that a variable has not yet been assigned to that name, so that cell was likely just not executed, or you aren't using the right variable name that you did assign
  - CAPITALIZATION and spaces matter
- ValueError: label "Acceleration" not in labels ('vehicle', 'year', 'msrp', 'acceleration', 'mpg', 'class')
  - When plotting, you may not refer to the column correctly. Above we see "Acceleration" should be lower-case to match the column name
- PatsyError: Error evaluating factor: NameError: name 'Acceleration' is not defined
  msrp ~ Acceleration
  - In the regression, again, typing the name of the column incorrectly will tell you it's not defined (in the table). Here again the capitalization is the cause.
- FileNotFoundError: File b'YOUR-FILE-NAME.csv' does not exist
  - The filename you gave doesn't exist. Make sure the file name is spelled *exactly* as it is in the directory.
- File "<ipython-input-15-1214ad7e0014>", line 1
  my_data = Table.read_table('Implicit-Disability_IAT.csv')
  SyntaxError: invalid character in identifier
  - If the students are copying and pasting, it's possible that you paste in the wrong

  '   **'**

  kind of apostrophe. There is a difference between   and   .

- Any "nan" in outputs, correlations or tables.
  - This is likely a result of merging tables, where there's no value for a corresponding column in the other dataset. There are many different methods of imputation, but to keep it simple, we've been just dropping observations with missing values:
    - new_table = Table.from_df(merged.to_df().dropna())
- TypeError: cannot perform reduce with flexible type
  - This *shouldn't* be happening anymore. But it essentially means that Python thinks that a particular column is a string (like a word) as opposed to a number. So it can't perform mathematical operations on words, and gives an error. If this happens, let us know and we can fix it in the dataset so that you don't have to work around this.