

Homework 01

[Re-submit Assignment](#)

Due Jan 26 by 12pm **Points** 100 **Submitting** a file upload

Available Jan 19 at 9:30pm - Apr 23 at 11:59pm 3 months

You must select 3 data sets from the 2020 in the TidyTuesday Github page. You must:

- Discuss the DATA DICTIONARY (if any) provided for the data set. State how many variables are discussed. Are the data types provided for the variables? Are there explanations? Would you feel comfortable accessing the data based on the DATA DICTIONARY?
- Read in the data set into a Pandas DataFrame. State how many variables are in the data set. What are the variable names? What are the data types? Are the variables consistent with the DATA DICTIONARY?

You must use a Jupyter notebook for the assignment. You must turn in your .ipynb source code and the rendered HTML file.

The TidyTuesday 2020 github page is: <https://github.com/rfordatascience/tidytuesday/tree/master/data/2020> [\(https://github.com/rfordatascience/tidytuesday/tree/master/data/2020\)](https://github.com/rfordatascience/tidytuesday/tree/master/data/2020)

If you CANNOT access the github repo, please let me know!