**Instructions for the ‘test.ipynb’ file**

* The file test.ipynb is a Python notebook file. It can be opened using Jupyter.
* You can install the Anaconda package from <https://www.anaconda.com/download/> and then install Jupyter within the Anaconda environment.
* Make sure to extract the contents of the zip file for our project in the working directory of your Jupyter. I will be usually the same as Anaconda.
* Once the zip is extracted, open the Jupyter app and then navigate to the path of test.ipynb. Open it.
* The results should already be displayed. If you wish to rerun it, just go to “Kernel > Restart and Run all” to rerun the code.
* Note: You may have to install packages in python/anaconda that are imported in the first 3 cells of the notebook before you run it.

**Instructions for ‘data cleaning.py’ file**

* The file ‘data cleaning.py’ is a python script file and can be opened using any Python IDE.
* The zip for our project is structured such that you will not have to move around the dataset files. The datasets and the cleaning file should be in the same folder.
* Set the working directory in your Python IDE to the directory of ‘data cleaning.py’ file.
* After that just Run the entire script to replicate the results.
* Note: You may have to install packages in python that are used in the script before you run it.