Meets Specifications

Dear Learner,  
I can see you did read the suggestions and comments from the previous review and followed the required guidelines to adjust your answer carefully.  
Excellent work!!!  
And I’m really glad that you paid so much patience to read the feedback and followed them to carefully amend the answers.  
You have a really great learning attitude, keep it UP!  
The work now meets all of the specifications. Congratulation. You made it!

**Please carry on your Great Work!**

**We must endorse the hard work and determination we perceived in this implementation and want this spirit to be maintained through out learning with us and we hope you will soon become one of us. Please keep up the spirit!**

If you encountered some challenges, please, feel free to share them with us. Thanks!

**Code Functionality and Readability**

All project code is contained in a Jupyter Notebook named wrangle\_act.ipynb and runs without errors.

Good work! All cells of the notebook run on my end without errors.

**Learning Notes**

I am a fan of using shortcuts with Jupyter Notebook. Check out [this medium post](https://medium.com/@venciityjordanov/jypyter-notebook-shortcuts-bf0101a98330) on **Jupyter Notebook Shortcuts**.

The Jupyter Notebook has an intuitive, easy-to-follow logical structure. The code uses comments effectively and is interspersed with Jupyter Notebook Markdown cells. The steps of the data wrangling process (i.e. gather, assess, and clean) are clearly identified with comments or Markdown cells, as well.

A remarkable job was done to organize the notebook in an easy-to-follow format.

**Learning Notes**

It is also good practice to use functions to avoid any code repetition.

* [Why use functions in programming](http://www.cs.utah.edu/%7Egermain/PPS/Topics/functions.html)?

Inline comments could also help with code follow up and ease the work of other programmers working on the same project.

* [Why should we comment code?](http://www.cs.utah.edu/%7Egermain/PPS/Topics/commenting.html)

**Gathering Data**

Data is successfully gathered:

* From at least the three (3) different sources on the Project Details page.
* In at least the three (3) different file formats on the Project Details page.

Each piece of data is imported into a separate pandas DataFrame at first.

Good work gathering data from three different sources and converting all of them into python data frames before performing wrangling.

Data was gathered from the following sources:

* Twitter-archive **CSV** file.
* Image prediction **TSV** file.
* **tweet\_json.txt** file.

**Assessing Data**

Two types of assessment are used:

* Visual assessment: each piece of gathered data is displayed in the Jupyter Notebook for visual assessment purposes. Once displayed, data can additionally be assessed in an external application (e.g. Excel, text editor).
* Programmatic assessment: pandas' functions and/or methods are used to assess the data.

Both visual and programmatic assessments are used in the notebook and the results are well documented.

At least eight (8) data quality issues and two (2) tidiness issues are detected, and include the issues to clean to satisfy the Project Motivation. Each issue is documented in one to a few sentences each.

Good work identifying quality and tidiness issues in the dataset.

**Cleaning Data**

The define, code, and test steps of the cleaning process are clearly documented.

The different steps of the cleaning process are clearly documented. We have the define, code and test steps which are clearly stated with some explanations of what process you intend to do at each level.

Copies of the original pieces of data are made prior to cleaning.

All issues identified in the assess phase are successfully cleaned (if possible) using Python and pandas, and include the cleaning tasks required to satisfy the Project Motivation.

A tidy master dataset (or datasets, if appropriate) with all pieces of gathered data is created.

Indeed, copies of the original pieces of the data are made before the cleaning process.

**Storing and Acting on Wrangled Data**

Students will save their gathered, assessed, and cleaned master dataset(s) to a CSV file or a SQLite database.

The cleaned dataset is saved to a csv file. Good work!

**Learning Notes**

* Check out [this StackOverflow Thread](https://stackoverflow.com/questions/16923281/pandas-writing-dataframe-to-csv-file) on **Pandas writing dataframe to CSV file**.
* Also, take a look at the [pandas.DataFrame.to\_sql](https://pandas.pydata.org/pandas-docs/stable/generated/pandas.DataFrame.to_sql.html) and this [Stackoverflow thread](https://review.udacity.com/Visualizations) for an example of saving pandas dataframe to SQLite.

The master dataset is analyzed using pandas or SQL in the Jupyter Notebook and at least three (3) separate insights are produced.

At least one (1) labeled visualization is produced in the Jupyter Notebook using Python’s plotting libraries or in Tableau.

Students must make it clear in their wrangling work that they assessed and cleaned (if necessary) the data upon which the analyses and visualizations are based.

Good work analyzing the cleaned data and plotting some visualizations of the data.

**Learning Notes**

* Check out the [python visualization documentation](https://pandas.pydata.org/pandas-docs/stable/visualization.html#visualization) for various ways of visualizing data.
* There are several other ways to visualize data including **Box plots, Line graphs, Pie charts. Check out** [**this documentation**](https://jakevdp.github.io/PythonDataScienceHandbook/04.14-visualization-with-seaborn.html) **on Visualization with Seaborn**.

**Report**

The student’s wrangling efforts are briefly described. This document (wrangle\_report.pdf or wrangle\_report.html) is concise and approximately 300-600 words in length.

The write-up is very detailed and within the limit required. Awesome job!

The three (3) or more insights the student found are communicated. At least one (1) visualization is included.

This document (act\_report.pdf or act\_report.html) is at least 250 words in length.

Great!

**Project Files**

The following files (with identical filenames) are included:

* wrangle\_act.ipynb
* wrangle\_report.pdf or wrangle\_report.html
* act\_report.pdf or act\_report.html

All dataset files are included, including the stored master dataset(s), with filenames and extensions as specified on the Project Submission page.