For this project, I thought that the wrangling part was a big tricky at first and required full understanding of the tasks at hand. I had trouble creating a Twitter developer account as Twitter simply replied account creation denial, even after I had taken the time to write well thought out paragraphs about the purposes with which I will be using their data. Therefore, I resorted to just downloaded the tweet json csv file provided by Udacity into my hard drive and loaded it onto Jupyter notebook from there. I was able to correctly request the image predictions tsv file and open it in Jupyter notebook. I also had no trouble with downloading the twitter archive.

After gathering all the necessary files, I begin accessing and cleaning them. After making a copy of the 3 datasets, the first quality tasks were to filter out tweets, or rows that were unnecessary. The project requires that all tweets used to be original tweets, so I filtered out rows that have a "retweet_status_id", then dropped all columns pertaining to retweeting and replying. Then, I filtered only tweets that have dog images attached to them. Afterwards, I dropped the 4 original dog stage columns of "doggo"," floofer", "pupper", and "puppo". For the next quality task, I changed filtered the unnecessary URL texts of the "source" column, and only the kept the values such as "Twitter for iPhone", the platform from which the tweet was created. The last quality task was to change the "rating_numerator" to float data type.

For the tidiness issues, I found 2 tasks that needed to be done. The first task was to combine the 'doggo', 'floofer', 'pupper', 'puppo' columns into a single "dog_stage" column, and if a tweet has more than 1 dog stage, I would concatenate the text together, separated with a comma. The second was task was to merge all 3 data sets by "tweet_id" as the primary key.

All of the issues that I have found are within the twitter archive file. While visually accessing the image predictions and tweet json files, I could not find any cleanliness or tidiness issues.