

**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.**

I had several problems gathering the data. In the first place, Twitter never replied to my request for developer access to their API, so I just had to use the code and .txt provided by Udacity. Secondly, I had a difficult time writing the code to read the JSON.txt file. My main conceptual problem here was just lack of familiarity with JSON; I did not realize that each JSON object was stored in the .txt file as a line. I found this code searching Udacity Knowledge and saw how it worked.

Assessing the data I found—I think—to be trivial (I probably missed many things). The only iteration that was required had to do with the `rating_numerator` and `rating_denominator` columns. I only examined the latter in the Assess section, and then found that I needed to look at the former as well. However, this only made sense to do after I'd cleaned the denominator issue and dropped the unusable denominators. So I wasn't sure whether to put this in the Assess or Clean section, and left it in Clean.

In the cleaning section, most of the tasks I set for myself were trivial and consisted mostly of recasting the `dtype` of some columns. The code that challenged me the most was required to create a breed column and sort through the most probably breed identification from the `p1_dog`, `p2_dog`, and `p3_dog` columns. This was helpful for me as I have problems with both `for` loops and `if` clauses in Python.

Finally, in terms of actually analyzing the data, I would have liked to make some more predictions using statistical models, but the only one I could think of was to try a multiple linear regression on the dog breeds, taking them as categorical variables, but since there are about 113, this seemed unduly complicated. I thought it would be useful to see if breed was a useful predictor of ratings. If the review has any feedback on this I would appreciate it.