

Udacity Act Report

[Menu](#)

Revised Wrangle Report

☐ Brian Campbell ☐ Uncategorized ☐ October 22, 2018October 25, 2018 ☐ 1 Minute

WRANGLE REPORT

Gathering Data:

Gathering data proved to be somewhat difficult. Figuring out how to use the Tweepy API took me 10 hours, and the project in its entirety took about 40 hours to complete. However, downloading the CSV and the TSV that Udacity provided was a straightforward task.

Assessing Data:

Assessing Data was a task of medium difficulty. In my initial assessments, I relied too heavily on visual assessments, and missed many errors. I reviewed my Data Analyst Nanodegree notes to identify programmatic methods of identifying errors. Once I implemented these methods, the data assessment became an easier and quicker task.

Cleaning Data:

The cleaning data process was by far the most difficult. However, the difficulty arose from my own carelessness. I did not take the time to organize the cleaning steps with the define, code, and test stages. I remedied that in the third submission of this project. Also, creating the stages column was laborious. I was unable to find a pandas method to help with the transformation, and I had to resort to 7 for loops. It was an effective, but not an efficient solution.

Additional Personal Insights from this Project:

I took two away three personal insights away from this project. First, the hardest and most time intensive part of data analysis is never the actual analysis; it is the data wrangling, assessing, and cleaning process. Second, clean data does not equate to statistically significant results. I found no causal relationships in my analysis, and my model explained a mere 0.2% of the variance of ratings scores. Third, Oxford is not the place for data analytics (I am in the Oxford MBA program). Oxford is ranked first in the global university rankings in the financial times, and the university still has not figured out how to connect to the internet. 10 of the 40 hours of my time can be contributed to internet connectivity problems. Do not believe the Financial Times!



Published by Brian Campbell

[View all posts by Brian Campbell](#)

[Blog at WordPress.com.](#)