

# wrangle\_report

## Project overview

This data wrangle and analyse project was very challenging. Even with well-designed project frame and detailed explanation, it took me quite a while to figure out how to start analysing this project. After I finished this project, I find it is very helpful to understanding the full processes of data analysing project. Analysing a full data set can develop into three main steps:

1. Data Wrangling
2. Storing, analysing, and visualizing wrangled data
3. Find the needed result and prepare report

I will go and explain what I did for all these three steps in details below.

## Data Wrangling

There are also three steps to do data wrangling: Gathering data, Assessing data and Cleaning data. Our main data sets are from three different resources: provided CSV document, download TSV document programmatic and download Json text file from API. This requires different set of data gathering skills to achieve getting those resources. Luckily these skills are all covered from classes and project details. The most challenging part is download Json file form API, which took me a lot time to figure out how to get the work done properly. In order to be a data analysis expert, the skills of gathering date from different resources are definitely needed to keep improving over the time.

Assessing date is the main part for me to understand all the data sets and find out the both quality and tidiness issues. My aim was to try get data set as clean as possible however it can hardly be done since there are so many addition issues appearing when I dig further. So, I focused on getting clean data for my insight questions alternatively.

It comes a slightly easier to clean the data once I find out the issues from assessing data step. I cleaned these data by using define, code, test steps. I found this is extremely useful that it can split one big mess data set into small issues.

## Storing, analysing, and visualizing wrangled data

Once I got a clean data set, it is a very import step to do is storing the clean data and use a copy of it to process analysis. Since this project is about rating dogs on twitter, I have listed a few interesting questions to me to process the analysis and have been able to find the answers to these questions with visualization tools. This is the most fascinating part to me to learn as a data analyst as it can give you a clear insight of relations between different data in a very powerful way.

## Find the needed result and prepare report

This step is mainly for preparing report to communicate with others and list the findings in order to solve certain problems.