Wrangling Data includes three parts:

1. Gathering Data
2. Assessing Data
3. Clean Data

Gathering Data

I gathered data from 4 resources:

1. image-predictions.tsv were already given and I downloaded manually
2. tweet_json.txt form tweepy library and I disappear my keys and I downloaded manually
3. twitter-archive-enhanced.csv were already given also and I downloaded manually
4 . some pictures from twitter page in the insights by HTML to help me in the insights and I downloaded programmatically

Assessing data

I assessed all the data both visual and programmatically

1. I used head() , tail() and sample() to the visual way
2. I used info() , isnull() and describe() to the programmatically

Quality issues

1.rename columns in image-predictions.tsv file
2.Some tweet have 2 different tweets_id, that are retweets
3. Missing values from images datasets
4. Delete the add numbers in timestamp in twitter-archive-enhanced.csv
5. Timestamp is string in twitter-archive-enhanced.csv
6. There some problem in the names of the columns so, I will make the names lower and replace '_' to one space
7. after marge there duplicate columns
8. In several columns null object are non-null (none to NAN) but its empty so I will drop it

Tidiness

1. Merge the three flies (image-predictions.tsv,twitter-archive-enhanced.csv,tweet_json.txt)
2. Merge rating numerator and rating denominator
3. doggo, floofer, pupper, puppo columns in twitter_archive_enhanced.csv should be combined into a single column as this is one variable that identify stage of dog.

Clean Data

1. Frist I make cope of data
2. Use pandas function to solve quality and tidiness issues
3. I used some functions like drop() , rename and other pandas functions to clean the data
4. Then I test my work to make sure everything is good