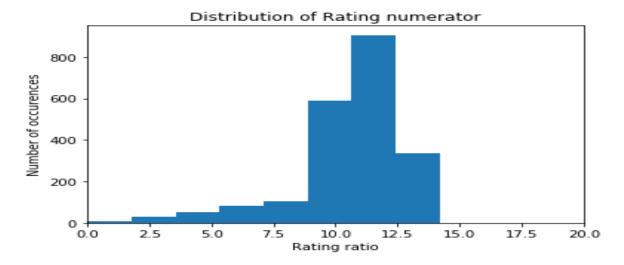# Act report

The master data frame which was created at the end was named "unified_df" has 2124 entries and 27 columns, the name of the columns and entries for the master data frame stated below:
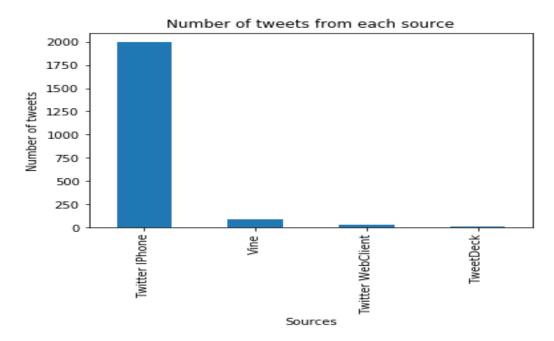
```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 2124 entries, 0 to 2123
Data columns (total 27 columns):
tweet_id                2124 non-null int64
in_reply_to_status_id   23 non-null float64
in_reply_to_user_id     23 non-null float64
timestamp               2124 non-null datetime64[ns]
source                  2124 non-null category
text                    2124 non-null object
expanded_urls           2124 non-null object
rating_numerator        2124 non-null int64
rating_denominator      2124 non-null int64
name                    2124 non-null object
dog_stage               356 non-null object
jpg_url                 2001 non-null object
img_num                 2001 non-null float64
p1                      2001 non-null object
p1_conf                 2001 non-null float64
p1_dog                  2001 non-null object
p2                      2001 non-null object
p2_conf                 2001 non-null float64
p2_dog                  2001 non-null object
p3                      2001 non-null object
p3_conf                 2001 non-null float64
p3_dog                  2001 non-null object
favorites               2124 non-null int64
retweets                2124 non-null int64
user_followers          2124 non-null int64
user_favourites         2124 non-null int64
date_time               2124 non-null object
dtypes: category(1), datetime64[ns](1), float64(6), int64(7), object(12)
memory usage: 450.3+ KB
```
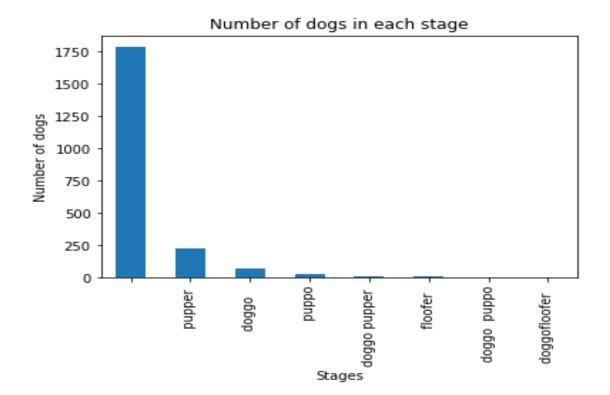
Visualization

First plot shows the distribution of rating numerator

The second plot shows the number of tweet from each source (Tweet For iphone, Vine, Twitter Web Client, and Tweet Deck)

The third plot show the number of dogs in each stage of the four stages (Pupper, doggo, Puppo, and Floofer)

Number of dogs in each stage

Insights

- Most of the rating were giving between 8 and 15 with some outlier ratings.
- Twitter for iphone is the main source, the other three sources are not so sourceful.
- Most of the dogs are in the Pupper stage and very few dogs are in the floofer stage. We have a huge number of dogs where their stage is not declared as represented in the plot with empty label, although there some entries have more than one stage.