

WE RATE DOGS - Data ANALYSIS

INTRODUCTION

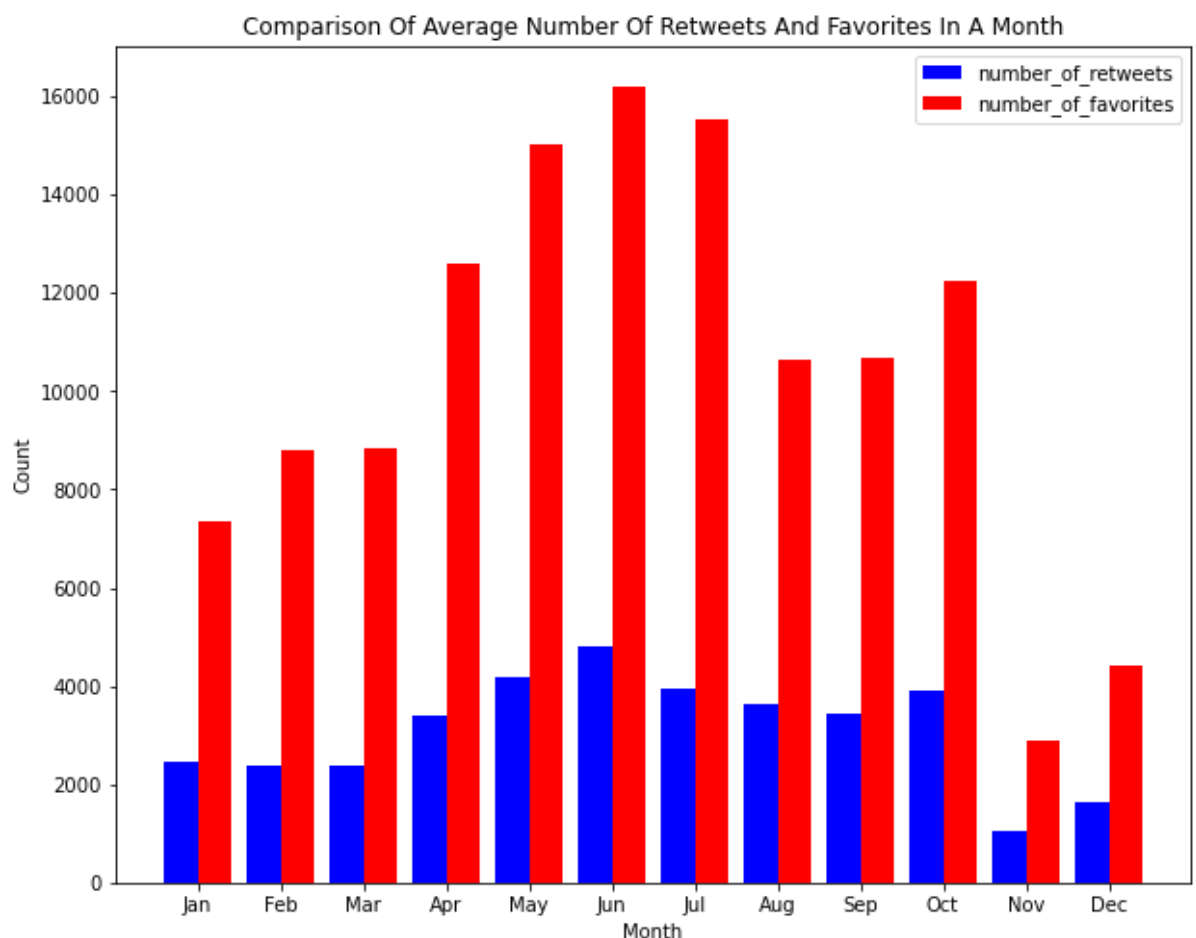
Here, I am going to summarize my analysis done in the project notebook. The dataset that I analysed is the tweet archive of Twitter user @dog_rates, also known as WeRateDogs.

WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. Why? Because "they're good dogs Brent." WeRateDogs has over 4 million followers and has received international media coverage.

ANALYSING THE DATA

Here, I have analysed the dataframes archive_clean, tweet_count_clean and image_pred_clean from the report after storing them into archive_master.csv file I have analysed the data on 2 different questions of interest.

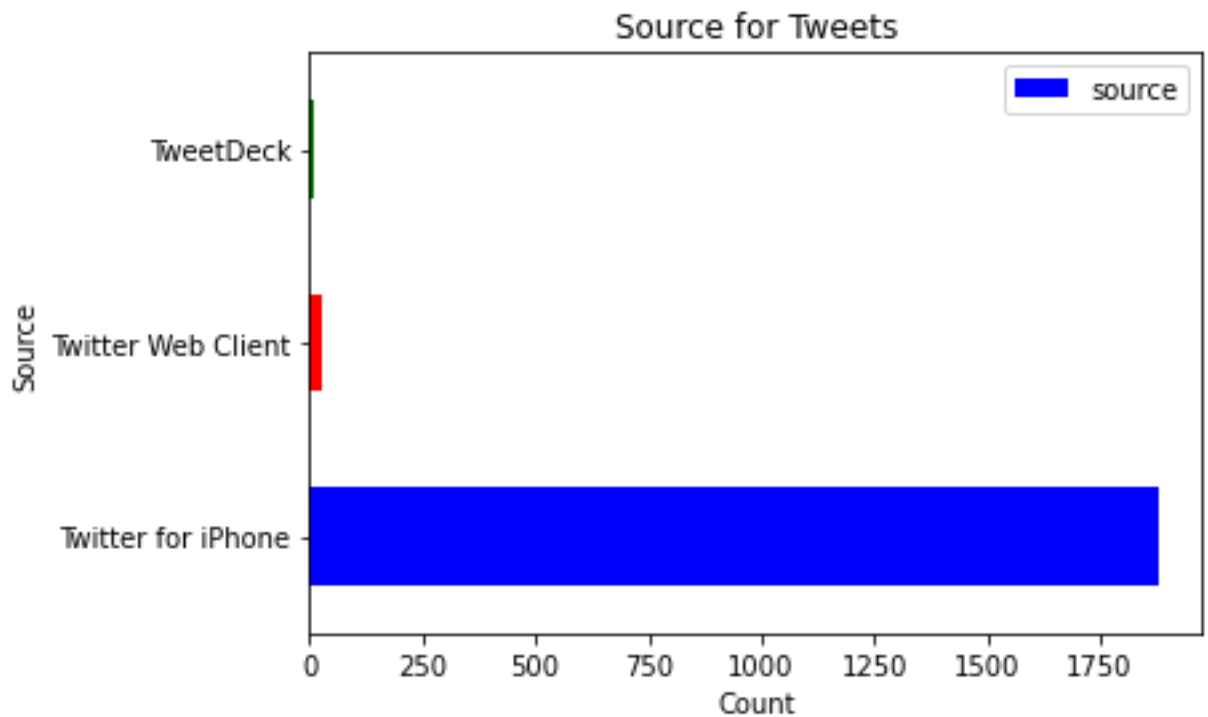
1. Comparison Of Average Number Of Retweets And Favorites In A Month



Observations:

1. As we can see that on an average, number of favourites are quite higher than number of retweets.
2. July has the highest average number of favourites and November has the lowest average number of favourites.
3. May has the highest average number of retweets and November has the lowest average number of retweets.
4. June has the second highest both average number of favourites and retweets.

Distribution of Source



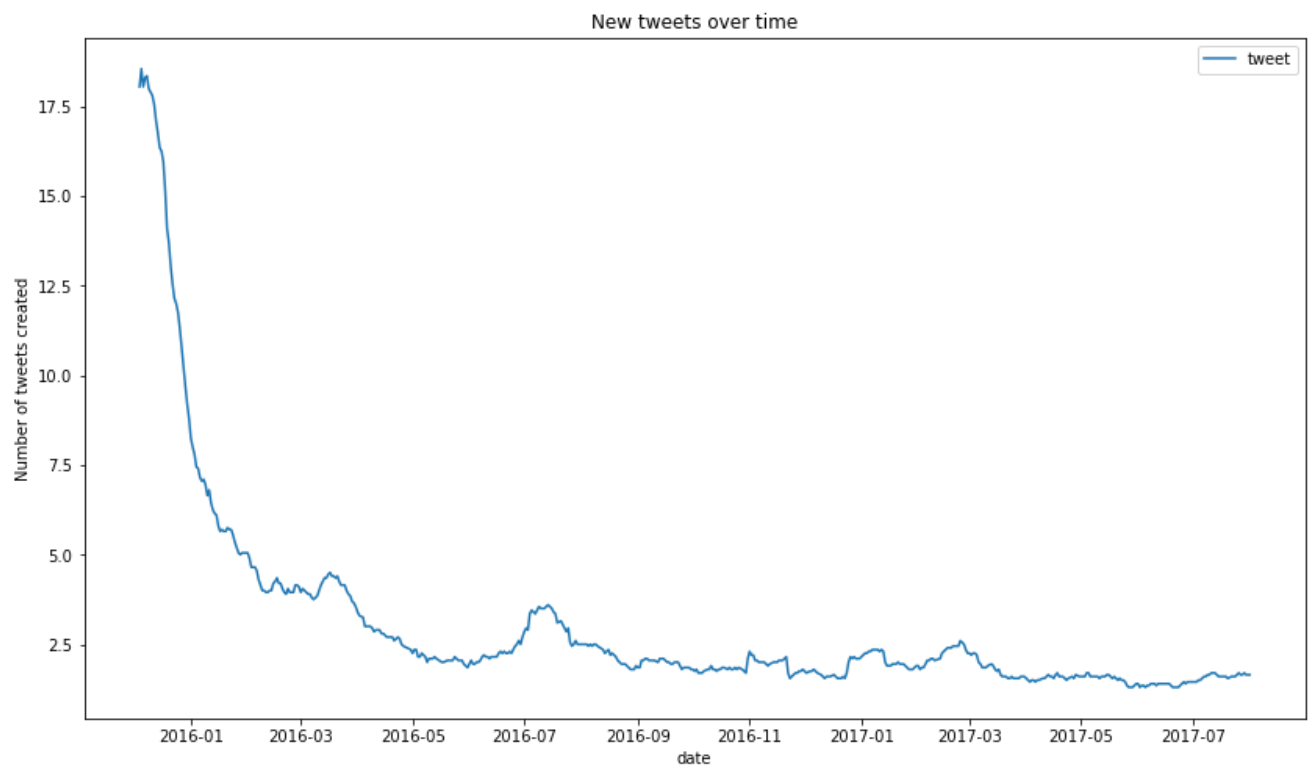
TwitterforiPhone 1882

TwitterWebClient 26

TweetDeck 10

Most of the tweets over 90% of them are sent via iPhone, next by Web and very few by TweetDeck

Tweets creation over time

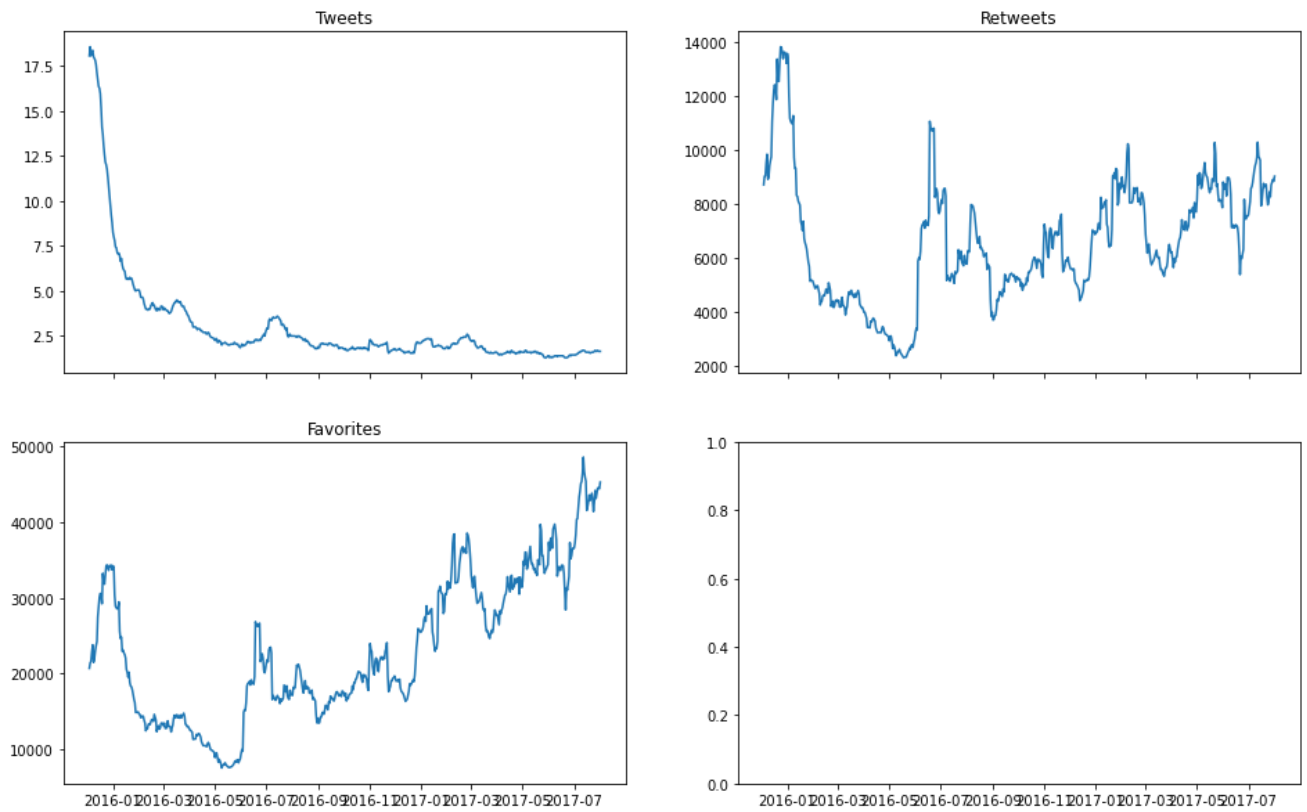


count 563.000000
mean 3.080728
std 2.995760
min 1.300000
25% 1.700000
50% 2.050000
75% 2.975000
max 18.550000

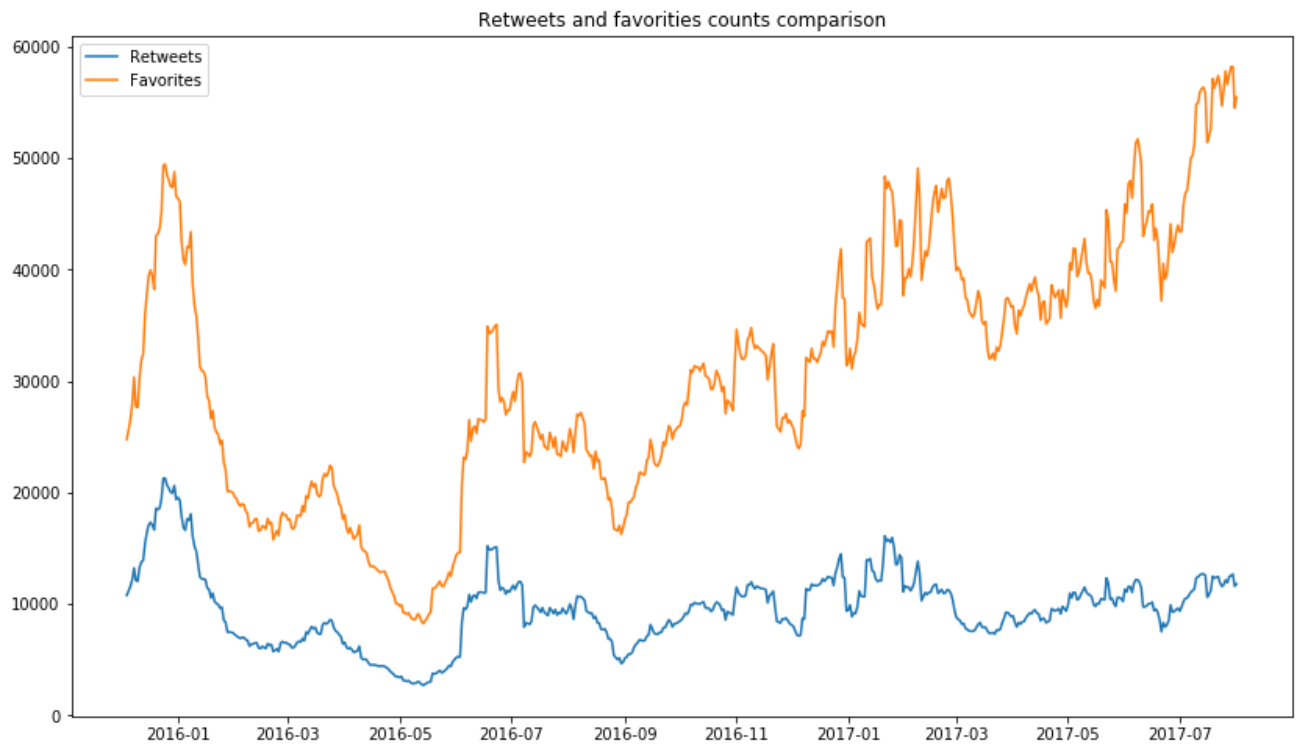
There were even more than 18 tweets on average created per day at the beginning of 2016 but this frequency quickly collapsed into about 3 on average in 2016 to be even lower in 2017. There is a clear trend of decreasing popularity in terms of creation of new tweets.

Retweets and favorities counts comparing to new tweets creation

Retweets and favorities counts comparing to new tweets creation



- Tweets - after initial boom, the amount of tweets created daily stabilizes on around 2.5 per day on average.
- Retweets - there is noticable variation of retweets but on average long-term they tend to stay on the same level.
- Favorites - after initial boom and then colapse, the trend is growing. It seems people still use historical tweets and marke them as favotes or the user base is growing which results in more and more favorites.



The number of retweets tends to stay at the same level whereas number of favorites grows over time.