Act Report

By Yasmeen Mubarak

10 Feb 2019

Analysis Of WeRateDogs

Real-world data rarely comes clean. Using Python and its libraries, we will gather data from a variety of sources and in a variety of formats, assess its quality and tidiness, then clean it. This is called data wrangling.

The dataset that we will be wrangling (and analyzing and visualizing) 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.

WeRateDogs downloaded their Twitter archive and sent it to Udacity via email exclusively for us to use in this project. This archive contains basic tweet data (tweet ID, timestamp, text, etc.) for all 5000+ of their tweets as they stood on August 1, 2017

After we wrangled the data we performed a basic analysis.

The Distribution Of Source

The plot shows that the most common Twitter users are from the iPhone app. Which means the most important source for Twitter is the iPhone app. Comparing with other sources, the users are very low.
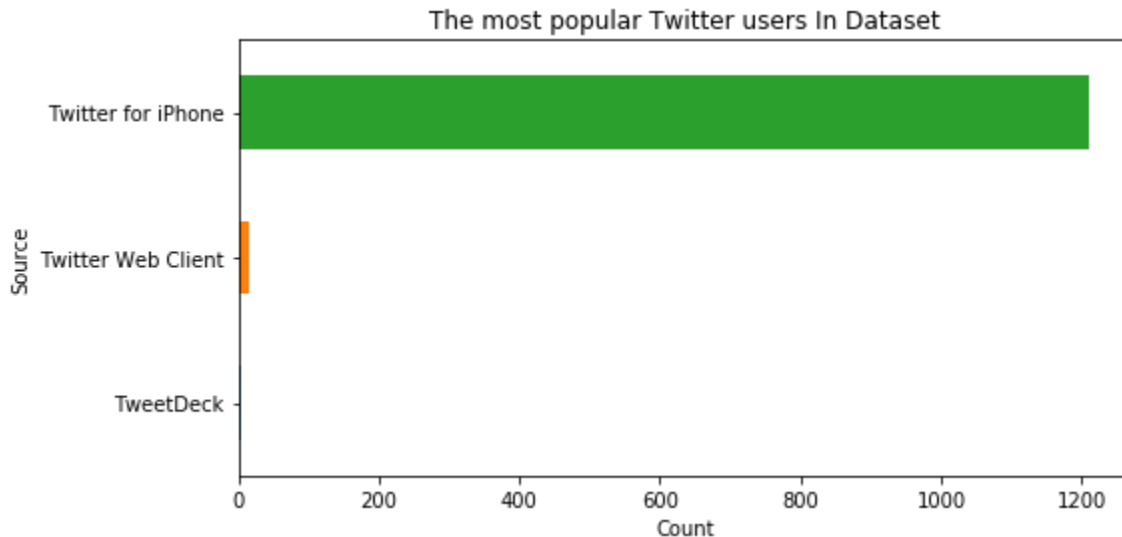
From the previous analysis we found that the Twitter app in the iPhone is the most used source, and based on this information we expect from Twitter to be very interested in ios users by develop the app continuously and caring for the ios design more.

Maybe the reasons that other sources have much lower users compared with iPhone app are:

- To view tweets from a certain list or only tweets containing your handle in the twitter web client, you have to navigate between views.

- Unable to schedule tweets in twitter web client.
- In twitter web client no tweets beyond 140 characters allowed.
- In twitter web client there is a delay between timeline updates.
- Scheduled tweets are in military time in tweetdeck.
- Tweetdeck program download required.
- Tweetdeck no built-in analytics feature.

if twitter want to attract more users into the other sources they need to work on these disadvantages.

# Relationship Between Favorite Count & Rating Numerator & Retweet Count

We have only one strong correlation between Retweet Count & Favorite Count, which mean the more tweet get retweet the more favorite will received.

We expected a relationship between Rating Numerator & Retweet Count and Rating Numerator & Favorite Count but from the correlation it's a weak correlation which means that as Rating Numerator increases or decreases, there is a lower likelihood of there being a relationship with the (Retweet Count & Favorite Count).Favorite and retweet counts are highly positively correlated as the correlation above shows.

It makes sense that if the tweet get more retweet obviously more people get to see it which leads to more people give the tweet a like.
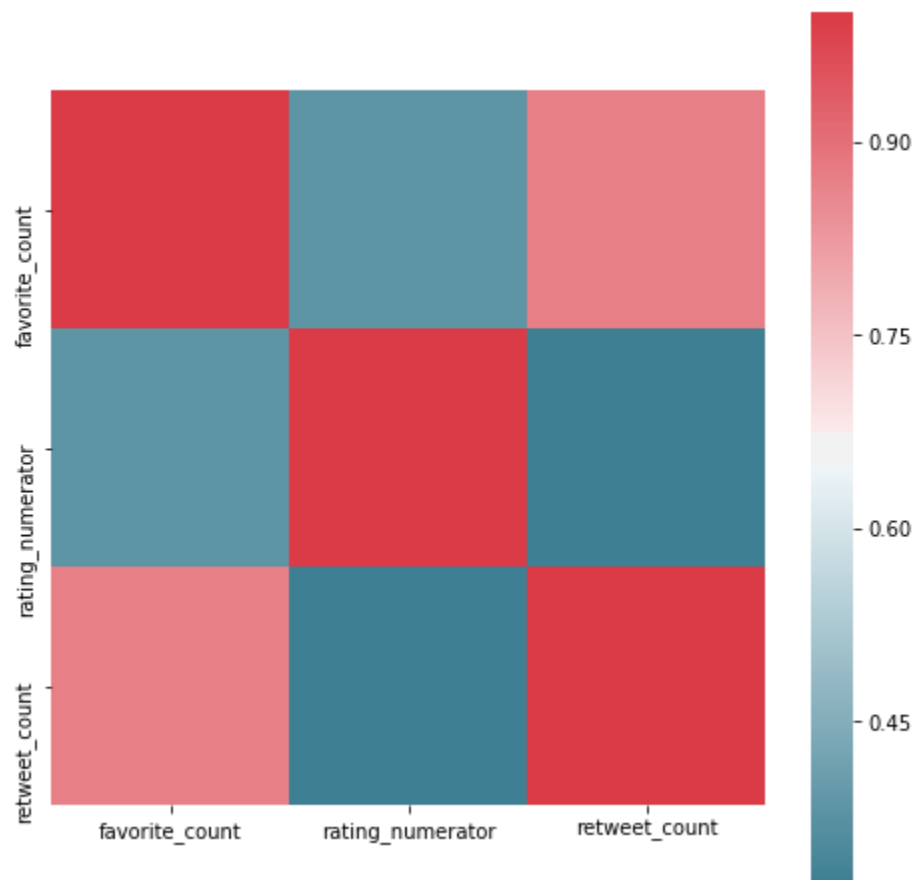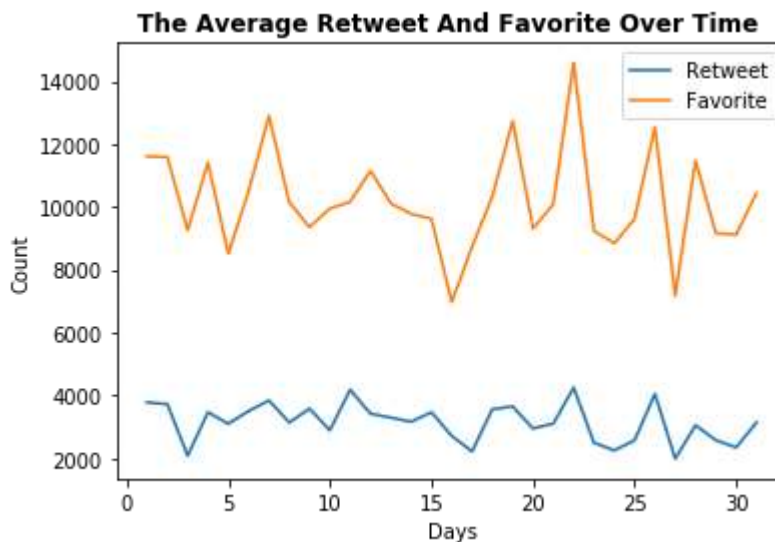
The Relationship Between Favorite Count & Retweet Count
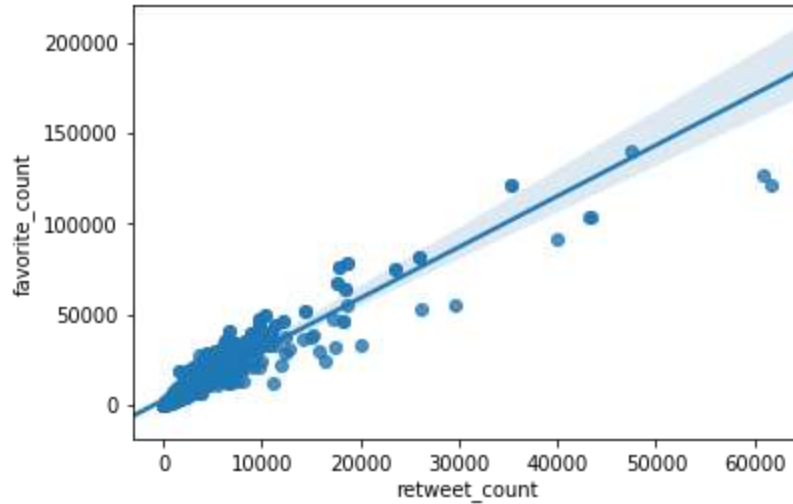
Favorite and retweet counts are highly positively correlated as the correlation above shows.

It makes sense that if the tweet get more retweet obviously more people get to see it which leads to more people give the tweet a like.

Retweet And Favorite Over Time

The Average Retweet And Favorite Over Time

The average of retweet and favorite is increasing and decreasing over time so we can't make conclusion about that, but there is a very clear conclusion the average number of likes much higher than retweet over time.

which make us wonder Why do so many people favourite tweets rather than retweet?

This question made us curious to find the reason and make us do a little search on the subject and we find some interesting answers :

- Bookmarking: some people use this feature to create reading lists.
- Attracting new followers: some Twitter users like posts of people whose attention they're interested to capture.
- Saying thanks: liking is quite commonly used to show gratitude.
- Testimonials: tweets of positive feedback or recommendations of their services/products to provide a handy resource for anyone looking for such comments.
- Keep track of favorite posts: there are those who simply like any particular tweet they enjoyed reading.

Obviously favorite have a lot of uses it's not just a simple like!

The Partitioning Of Doge Stage

The chart shows that people tend more to acquire pupper dogs maybe because they are small and so cute. And the ratings shows the doggo and puppo stage had the highest rate. for sure there pictures were very nice and lovely.

So puppers are getting much lower rates than the other dog types. We think the lower rating because pupper can be irresponsible dogs.

Floofers the lower rate they get is above 10. Maybe because they are so cute all the time.

Repartition Of Dog Stage