

act_report

August 20, 2019

0.1 Data Wrangling Project: WeRateDogs

Project Introduction: The dataset which is going to be wrangled 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. WeRateDogs has over 4 million followers and has received international media coverage.

In this report, you may find some conclusions and insights based on given datasets:

1. twitter-archive-enhanced.csv: The WeRateDogs Twitter archive. **2. image-predictions.tsv:** Predictions for breed of dog according to a neural network. **3. The Twitter API & JSON:** Each tweet's retweet count and favorite ("like") count.

Research questions: 1. Which dog breeds are popular? Find top-10! 2. Is there a time-trend of rating? 3. Is there a correlation between favorites, ratings and number of retweets? How strong is it? 4. Try to find the proportion of genders of dogs. 5. Find the photo of the most favoured dog!

Top-10 Dog Breeds Here, breeds were compared based on their popularities. In this study, 'popularity' means total number of retweets and favorites!

Conclusion: Irish_water_spaniel seems the most popular breed! Boston bull and dalmatian seem pretty close to each other.

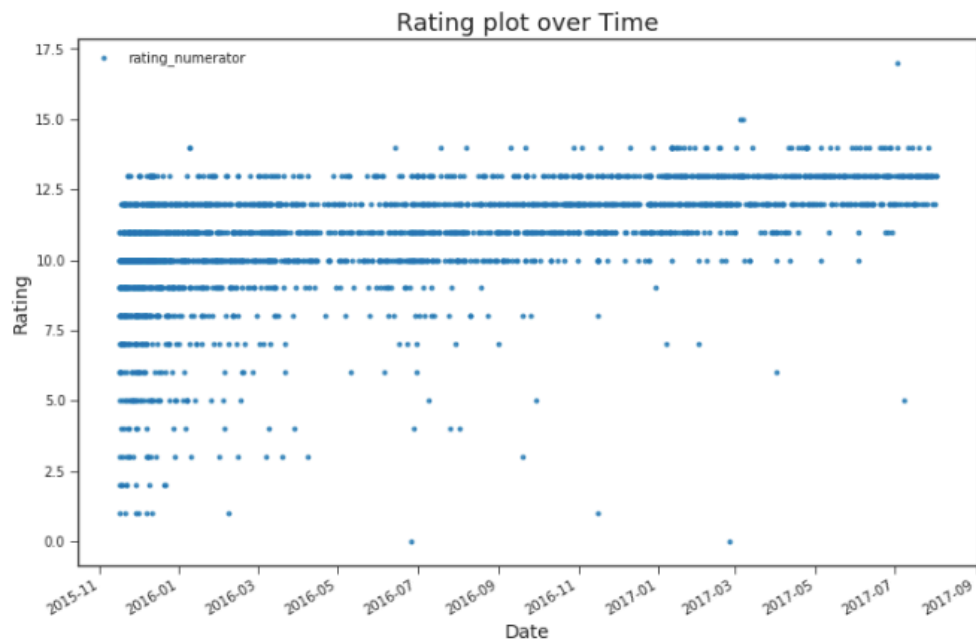
Time Trend of Ratings! A plot is given which contain ratings and their corresponding dates so as to shown the trend.

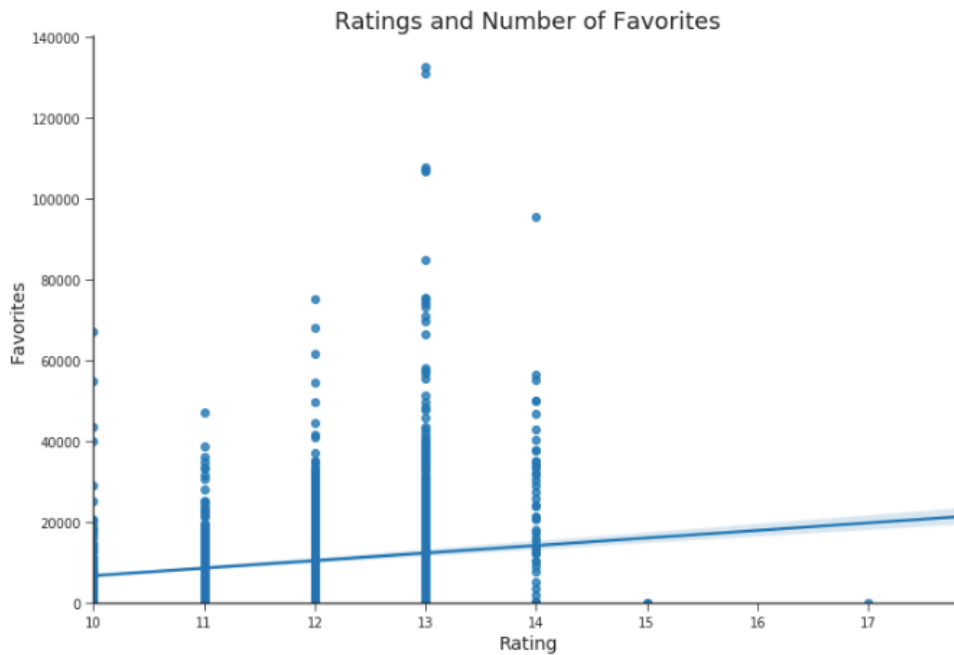
Conclusion: Above diagram shows that ratings increase over time!

Correlations Between Number of Favorites, Ratings and Number of Retweets!

Intuitively, we expect a pairwise correlation rating and between number of favorites. Because we give high ratings to our favorites! Let's check the correlation between rating and favorites!

	popularity
prediction	
Irish_water_spaniel	28617.000000
Boston_bull	25396.476190
dalmatian	25039.857143
Norwich_terrier	24733.600000
Bedlington_terrier	24495.800000
Rhodesian_ridgeback	22608.714286
Bouvier_des_Flandres	22511.000000
Mexican_hairless	22235.833333
Pomeranian	20785.717391
Labrador_retriever	20227.737288

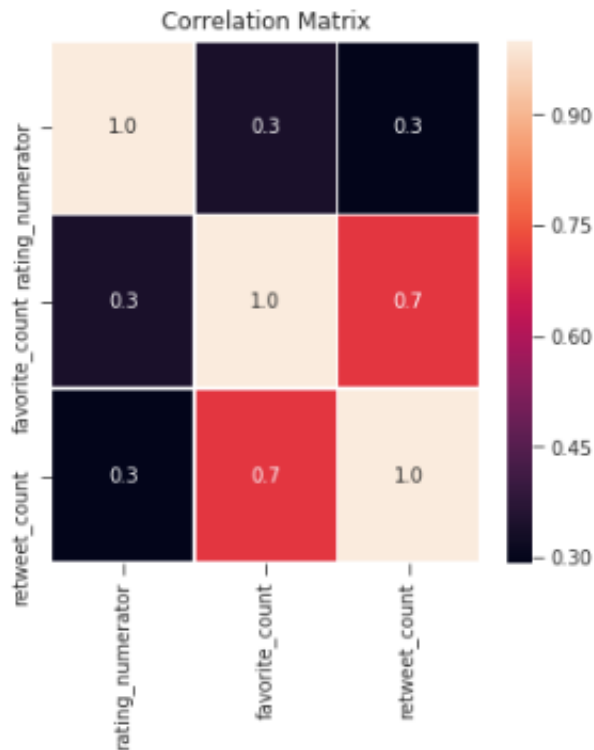
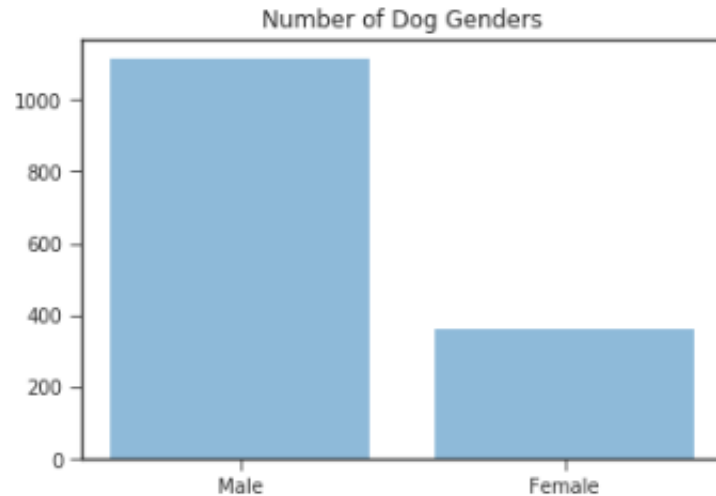




It's hard to see whether there is a correlation between ratings and favorites. Let's calculate the pearson's correlation coefficient.

	favorite_count	retweet_count
favorite_count	1	0.702666
retweet_count	0.702666	1

There is a moderate correlation (~70%) between number of favorites and number of retweets. Now, let's also consider the retweets.



Conclusion: There is a moderate correlation between number of favorites and number of retweets. However, those two have a weak correlation with ratings!

Proportion of Dog Genders!

This information is not given directly however it's basically derived from the tweets based on some words that describes or signs a gender such as he, she, boy, girl etc.

Conclusion: It's clear that number of male dogs are significantly higher than female dogs.

Ladies and Gentlemen! Here is the most favored dog: But please keep it mind that 'they're all good dogs :)'



0.1.1 Summary

Irish_water_spaniel is the most popular breed currently. Good news is there is increasing trend for the dog ratings and this can be considered as that people are getting used to love dogs more. However, this is just an insight. There might be other factors such as increased number of tweeter users or people are getting to spend more time in twitter etc.

Based on the data derived from the WeRateDogs tweeter account, people who retweet a dog also make it his/her favorite likely. However, no significant correlation found between ratings and number of favorites (and between ratings and number of retweets as well!). It seem odd initially because intuitively we might expect that most favorited and/or most retweeted dogs should also have high ratings. However it is not impossible. Think about it, in almost every social media platform (such as twitter, instagram and facebook) there are 'haters'! And these people are tend to give dislikes, low ratings etc. however for some reason they might retweet the tweets as well for example to say "see this ugly dog"! However, they are just 'haters'! The fact that 'they're all good dogs :)'.

Interestingly, there are much more male gender dogs than female gender dogs. Most probably they are easy to take care however it's just a guess, there is no knowledge derived from the given datasets.

Reported by Ugur URESIN, MSc. Thanks!