Analysing & Visualising @WeRateDogs

By Bayan AlArifi

WeRateDogs™ ✔

@dog_rates

Your Only Source For Professional Dog Ratings   IG, FB, Snapchat ⇨ WeRateDogs partnerships@weratedogs.com

⊙ DM YOUR DOGS

🔗 weratedogs.com

📅 Joined November 2015

Tweet to   Message

Introduction

This reports on the analysis and visualization of the @WeRateDogs tweets. This Twitter account rates dogs with humorous commentary. The rating denominator is usually 10. However, the numerators are usually greater than 10. This aspect was not cleaned as it is part of the humor and popularity of WeRateDogs. I explore the account's popularity over time, the top tweet in terms of number of retweets or favorite, the top dog breed, and top dog stages. On average, the @WeRateDogs tweet length is 120 characters.

Data

The data was gathered from multiple sources, as listed below. The twitter archive is great, but it only contains very basic tweet information. Additional gathering, then assessing and cleaning was required for "Wow!"-worthy analyses and visualizations.

- **Load @WeRateDogs twitter archive file**

  Load the **@WeRateDogs** tweets archive (2356 tweets). This archive contains basic tweet data (tweet ID, timestamp, text, etc.) for all 5000+ of their tweets as they stood on August 1, 2017.

- **Extracting a file from a server**

  The tweet image predictions, i.e., what breed of dog (or other object, animal, etc.) is present in each tweet according to a neural network. This file (image_predictions.tsv) is hosted on Udacity's servers and is downloaded programmatically using the Requests library.

- **Extract data from twitter API**

  Retweet count and favorite count (likes) are two of the notable column omissions from the Twitter archive. Fortunately, this additional data can be gathered, along with any additional interesting data by using tweepy library. The tweet IDs in the @WeRateDogs Twitter archive were used to query the Twitter API for each tweet's JSON data, and store each tweet's entire set of JSON data in a file called tweet_json.txt.

Basic Stats

- The @WeRateDogs account got between 12 and 61,934 retweets for its original tweets.
- The @WeRateDogs account got between 78 and 140,343 favorites for its original tweets.
- The neural network performed the best on the 1st iteration with a mean prediction of 0.6.
- The average rating numerator given was 11.
- The least rating numerator given was 0, and the most was 144.

Most and Least Popular Tweets

- **The most favorited tweet is:** Here's a super supportive puppo participating in the Toronto #WomensMarch today. 13/10 https://t.co/nTz3FtorBc **No. of favorites:** 140,343

WeRateDogs™ ✔
@dog_rates

Follow

Here's a super supportive puppo participating in the Toronto #WomensMarch today. 13/10

10:26 AM - 21 Jan 2017

47,678 Retweets 140,318 Likes

Mary-Margaret Jones

689   48K   140K

- **The most retweeted tweet is:**
  Here's a doggo blowing
  bubbles. It's downright
  legendary. 13/10 would watch
  on repeat forever (vid by Kent
  Duryee)
  https://t.co/YcXgHfp1EC

  No. of retweets: 61,934

WeRateDogs™ ✓
@dog_rates

Follow

Here's a doggo blowing bubbles. It's
downright legendary. 13/10 would watch on
repeat forever (vid by Kent Duryee)

0:19 4.42M views

4:31 PM - 4 Jun 2016

61,918 Retweets 121,279 Likes

951   62K   121K

- **The least favorited and least
  retweeted tweet is:**
  Oh my. Here you are seeing an
  Adobe Setter giving birth to twins!!!
  The world is an amazing place. 11/10
  https://t.co/11LvqN4WLq

  No. of favorites: 78
  No. of retweets: 12

WeRateDogs™ ✓
@dog_rates

Follow

Oh my. Here you are seeing an Adobe Setter
giving birth to twins!!! The world is an
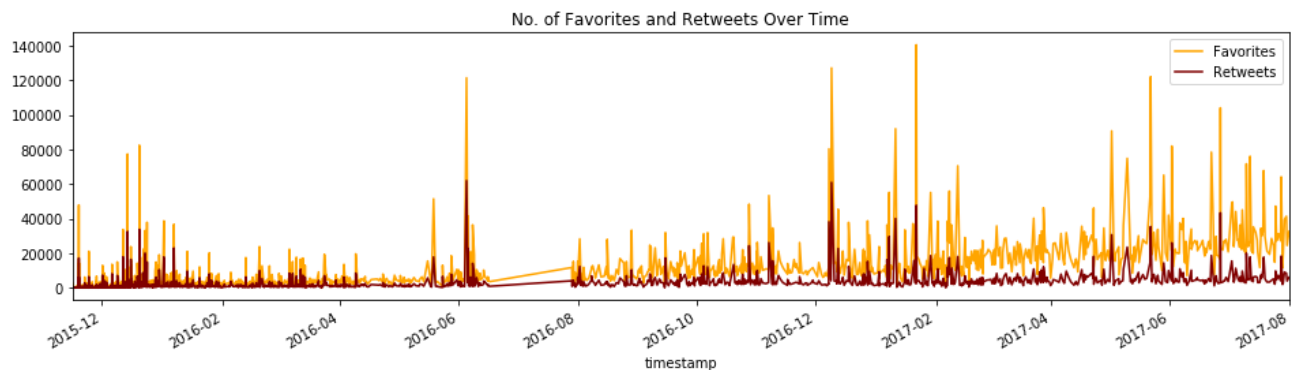amazing place. 11/10

7:55 PM - 15 Nov 2015

12 Retweets 78 Likes

12   78

3

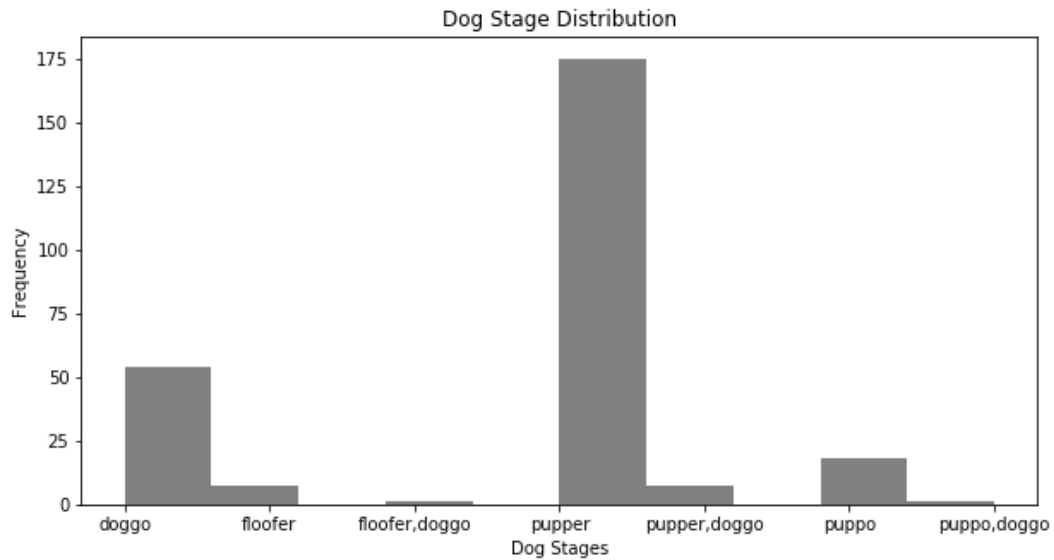# Number of Retweets and Favorites Comparison

I compared between the number of retweets and favorites that the @WeRateDogs account has over time. I concluded that;

- We can see from the plot that in general, @WeRateDogs tweets get more favorites than retweets.
- In June 2016, the account got its highest number of favorites (140,343 favorites), and its highest number of retweets (61,934 retweets).
- In March 2016, the account got its lowest number of favorites (78 favorites), and its lowest number of retweets (12 retweets).
- In general, there's a pattern where the favoriting and retweeting behaviour are similar in their rises and falls, where the favorite_count is usually double the retweet_count. Except in March 2017 going forward, the favorites counts started to exceed the retweets counts by far.
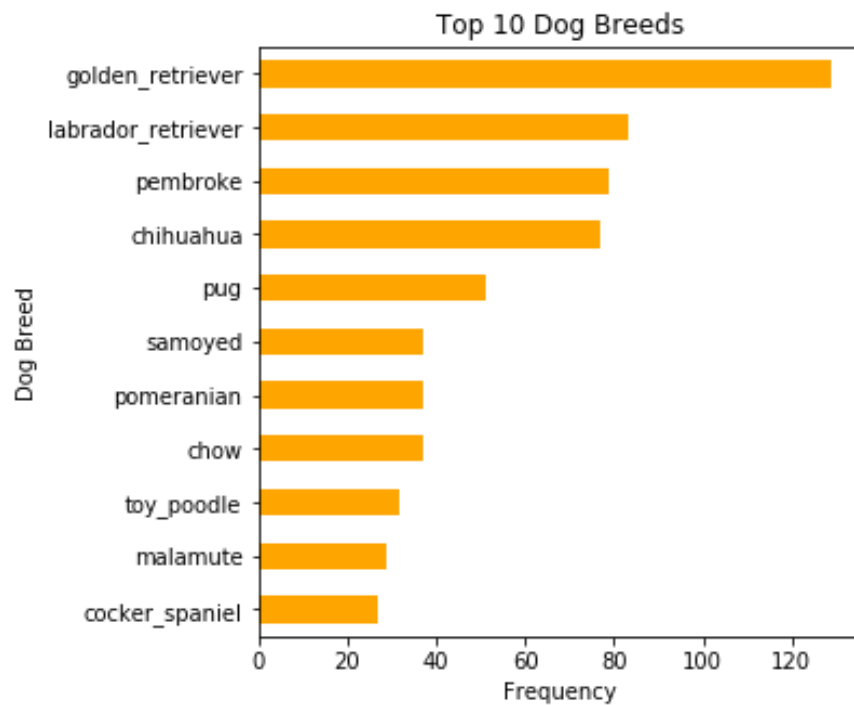

No. of Favorites and Retweets Over Time

Most Popular Dog Stage

Pupper is the most popular dog stage with a total of 175, and floofer is the least popular. However, most tweets have no dog stage specified and were excluded. The plot below shows the dog stages distribution.

Dog Stage Distribution

Most Popular Dog Breed

An image prediction algorithm was used to predict the dog breed from the images. It has found the golden retriever as the most common dog breed, with a total of 33. The plot below shows the top 10 predicted dog breeds.

Top 10 Dog Breeds

Basic Sentiment Analysis:

It doesn't come as a surprise that most @WeRateDogs tweets are positive since the whole account is created to express dogs love and appreciation.

Percentage of positive tweets: 51.47%
Percentage of neutral tweets: 27.02%
Percentage de negative tweets: 21.51%