Analyzing and Visualizing Data for WeRateDogs Twitter

by Alena Sukretna

Main areas that I explored on my wrangled data set:

1. Dog ratings
2. Favorites
3. Retweets
4. Dog breed prediction(p1)

1. Dog ratings

Maximum dog rating is 1776 out of 10 (for Atticus) and this is not a typical rating.
Picture of Atticus:

Average rating is 12.24 out of 10. This happens because people like dogs and overrated them on purpouse.

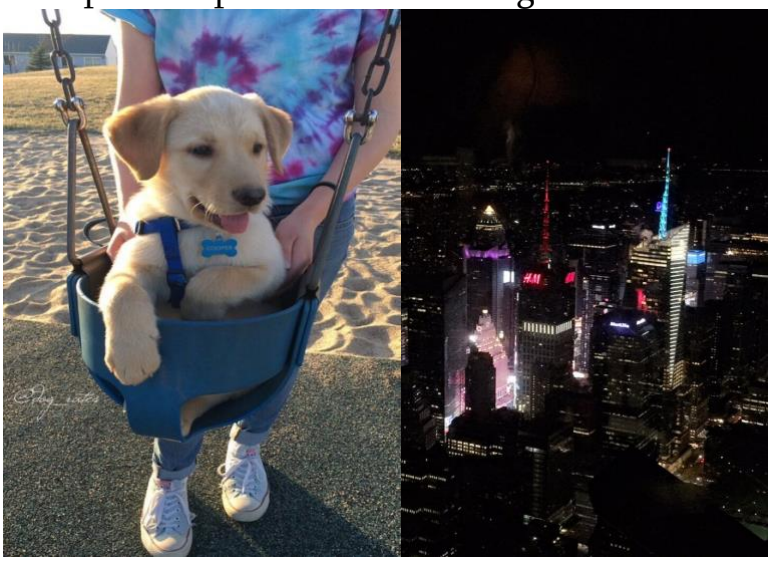
Is the number of retweets depends on dog rating?

As correlation shows there is not much connection between ratings and retweets.

Let's look on **unrated dogs**:

Surprisingly, underrated dogs have high number of retweets and favorites.

Let's pull the pictures of these dogs to see the reasons of low ratings



So, we have: dog in kids swing wich has 24183! favorites and Ney York at night and almost invisible black dog.

2. Favorites (likes)

The maximum count of favorites(144,108) has puppo without name which "march for my moms".

The average number of favorites is 8900, minimum 80.
Correlation shows that there is not much connection between ratings and favorites either.

3. Retweets

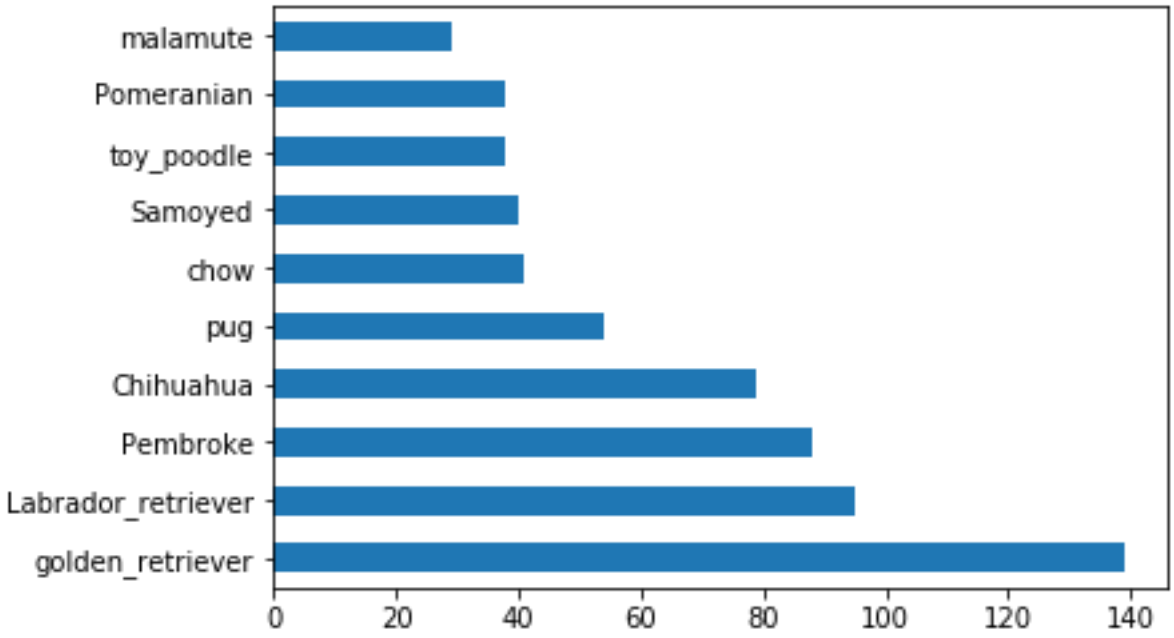
There is a strong correlation between favorites and retweets. As a tweet becomes Favorite, it's retweets increase.

Maximum retweets (77944) however has another dog without name with a message:"Here's a doggo realizing you can stand in a pool."

4. Dog breeds

The most popular predicted dog breed is golden retriever with mean confidency 0.72 and all images detected as dogs. There is some non dog
related predictions like laptop, flamingo, orange, etc.

This bar plot shows 10 the most popular predicted dog breeds:


malamute
Pomeranian
toy_poodle
Samoyed
chow
pug
Chihuahua
Pembroke
Labrador_retriever
golden_retriever
0
20
40
60
80
100
120
140