

@WeRateDog

Analyzing and Visualization Data

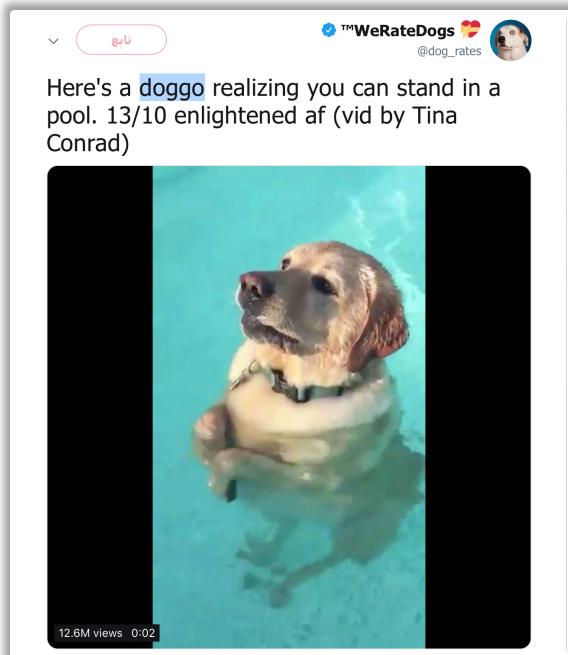


Introduction

@WeRateDogs is famous tweeter account for share images for cute and funny dogs. The account is do some rates those dogs and write funny comments on the photo of people dogs we collect some from multiple data source and I do some cleaning and wrangling methods and also use visualization and analysis tools like matplotlib to answer some wondering questions like whom famous dog !!

1. Top Five Favorite Dog





1. Doggo

He is got more than 164005 likes from users is I think he is famous dogs.

2. Puppo

This second dog in our list his got 140166 likes from Users.



3. Stephan

His 3rd famous liked dog in our list his rate 13/10 and he got more than 126822 likes.

4. Jamesy

This is 4th dog in our list it is got 121776 likes and his rate 13/10

↙ ↘

TM WeRateDogs ❤️ @dog_rates

This is Jamesy. He gives a kiss to every other pupper he sees on his walk. 13/10 such passion, much tender



٢٠١٧ - ٢١ مارس ٥:٣٨

٦٥,٣٥٨ إعجابات | ١٢١,٧٧١ تعليقات

↙ ↘

TM WeRateDogs ❤️ @dog_rates

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



4.42M views 0:17

٢٠١٦ - ٣٤ مارس ٤:٣٣

٦١,٧٩٥ إعجابات | ١٢١,١٥٣ تعليقات

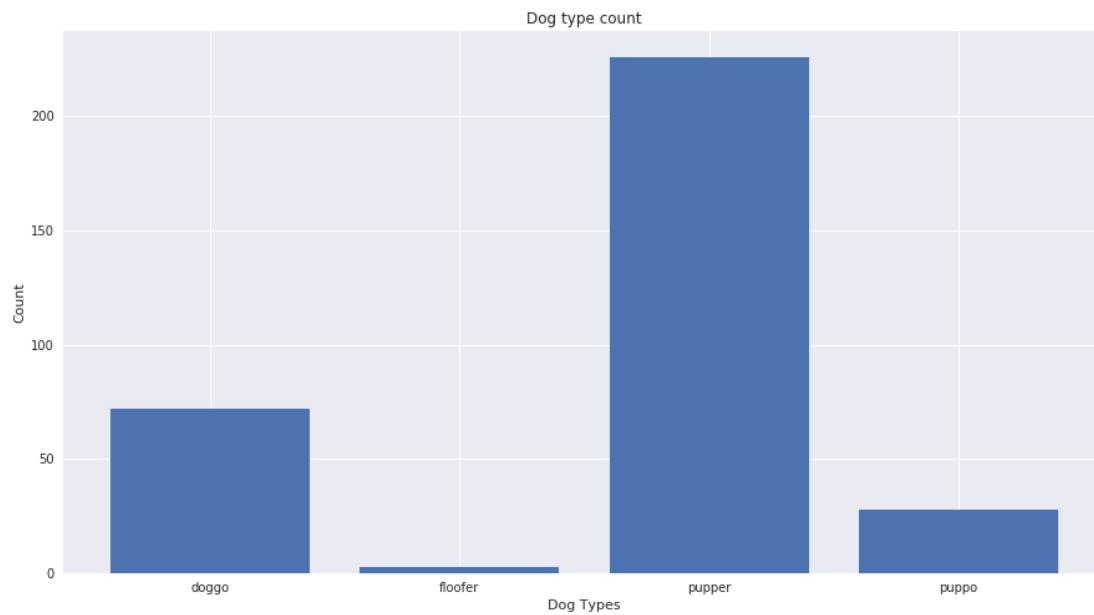
ألف ١٢١ إعجاب ٦٢ ٩٥

5. Doggo

This last dog in our top 5 his favorite count is 121151

2. What is most popular Dog Kind over all dataset?

I used histogram to find what is popular kind between all dog.

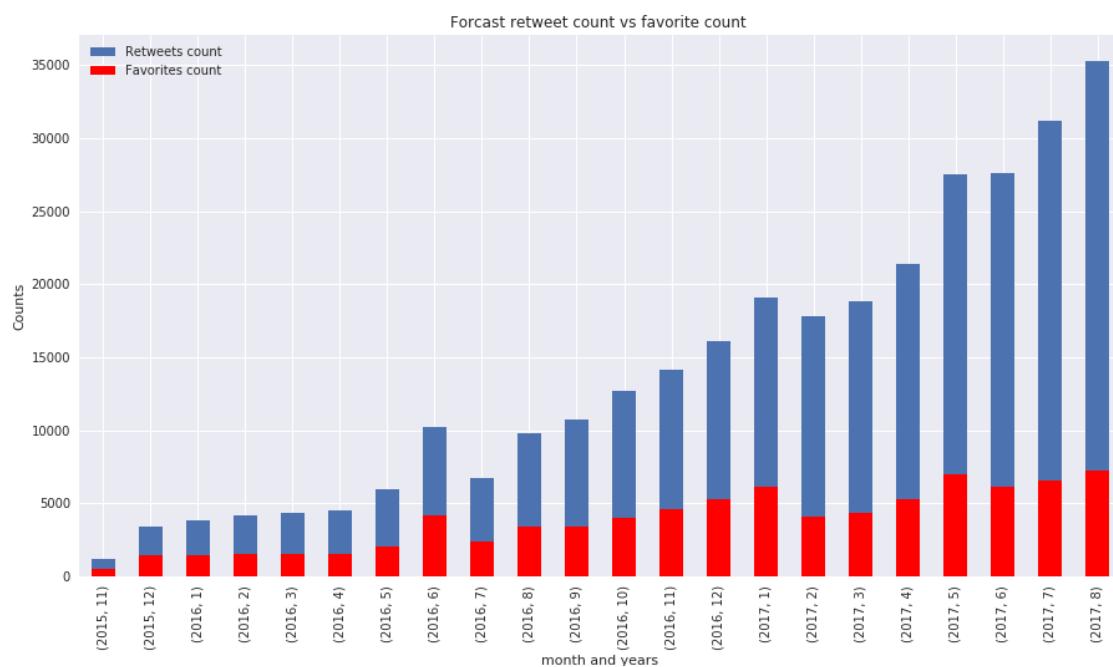


dog_type	
doggo	72
floofer	3
pupper	226
puppo	28

The Popular dog is pupper is comes 226 times over the time.

3. What is Forecast or time counts between the favorite and retweet over year?

In this section I try to compare count of retweet vs favorite for each months of years and I find the trends is goes up and positive and increase each month for retweet while there A multimodal distribution.



4. What is Relation between favorite count and retweet?

Using scatterplot to figure out the relation kind between the favorite and retweet count the plots shows there is high positive correlation between variables except some of values it shows as outliers and the correlation value between them is 0.9296

