

Analysis and Visualization of Data

Three (3) main insights were derived from the dataset, including the following:

1. Dog stage with the highest average ratings
2. Top 5 tweets, by the number of retweet_count
3. Top 5 favorite dog_breeds by the first prediction.

Pupper dogs came out as the dog with the highest average ratings of about 12.36, followed by the doggo dogs with an average rating of 11.95. Due to the fact that there are tweets with more than one dog, i.e., a tweet with both a doggo and floofer or doggo and pupper, or doggo and puppo, all tweets that fell into this category had the same average ratings of 10, all falling into the least rating. The dog stages with the lowest average rating were the puppo and floofer dogs.

Looking at the top 5 tweets with the highest no of retweets counts, the doggo dog appeared twice, with a doggo dog coming out as the highest retweeted dog, with tweet id 744234799360020481, tweeted on the 18th of June 2016, with 79515 no of retweets, talking about “the dog realizing you can stand in a pool”. It can also be observed that the lowest no of retweet was 0 with tweet id 838085839343206401 tweeted on the 4th of March 2017. It can also be noted that a puppo dog came out among the bottom retweeted dog tweets with just 3 retweets.

On the account of top favourite dog breeds by the first dog prediction, a puppo dog of the Lakeland_terrier breed with tweet id 822872901745569793 came out as the highest favorite count of 132810, tweeted on the 21st of January 2017, though not the highest retweeted dog, but among the top 5.

In visualizing the retweet_counts over time (year and month), a line chart was used. From it, it can be deduced that more retweets were made in 2016, with the least number of total retweets in 2015. Also, more retweets were made in December, and the least being August.

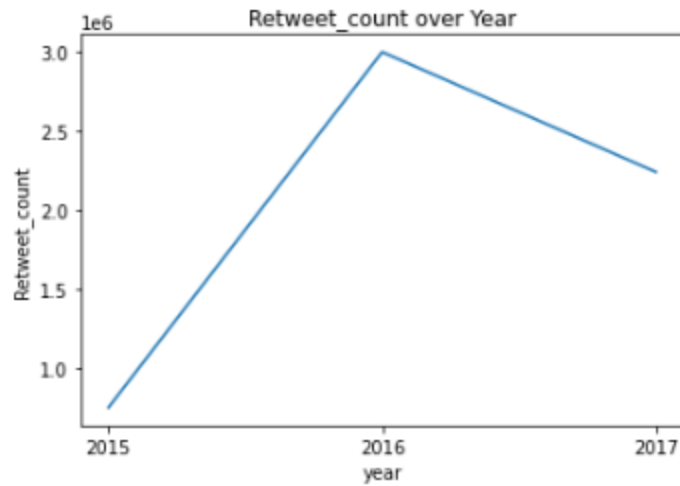


Figure showing retweet_count over year

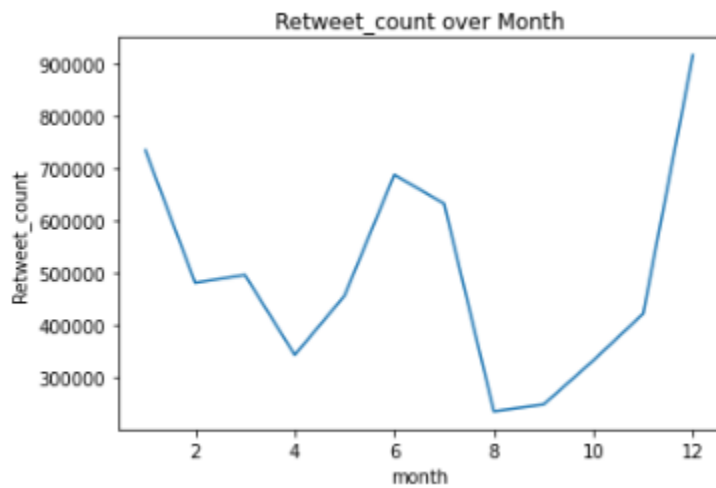


Figure showing retweet_count over month

More robust insights were visualized, to get the particular images relating to the dog with the highest number of retweets and dog with the highest number of favorite_counts.



Figure showing the dog with the highest no of retweets

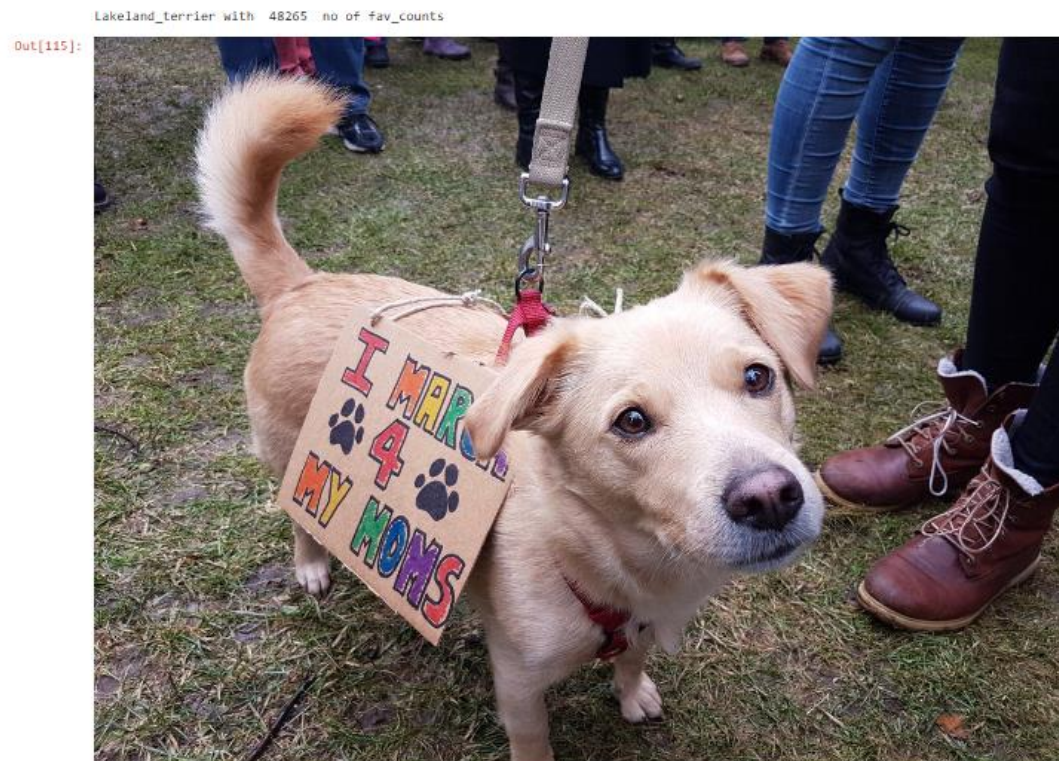


Figure showing the dog with the highest number of favorite counts.