

act_report

February 15, 2022

0.1 Report: act_report

- Create a **250-word-minimum written report** called "act_report.pdf" or "act_report.html" that communicates the insights and displays the visualization(s) produced from your wrangled data. This is to be framed as an external document, like a blog post or magazine article, for example.

1 WeRateDogs twitter data wrangling

A favorite account on Twitter, of many, including myself, is WeRateDogs. This account has been charming the masses with adorable tweets about all the good dogs that people send them to rate. The ratings are consistently better than a 10/10, as everyone knows, dogs are perfect and we don't deserve them.

In order to gain some insights regarding their Twitter archive, I used multiple sources of information. The one with the greatest amount of information was the Twitter archive. Information regarding favorites and retweets was also needed. This was acquired via the API. Last was a collection of images pulled from the feed.

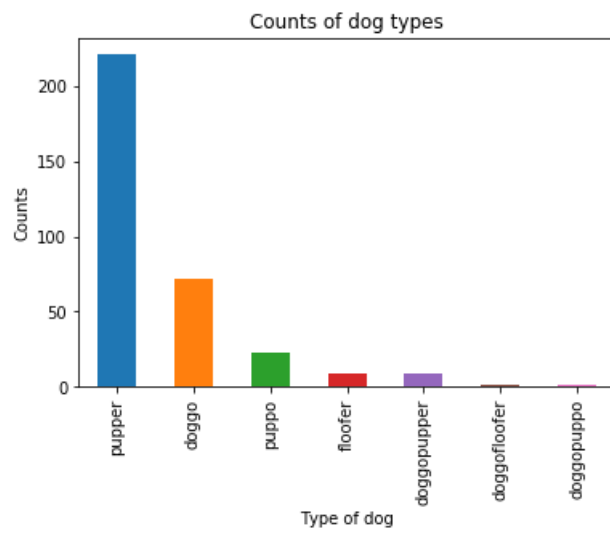
After the information had been gathered, a few things had to be corrected. One of the datatypes, specifically the one involving the date was incorrect. There were also multiple unnecessary columns that needed to be addressed. Unneeded columns were dropped, some columns were combined, and invalid data was corrected. After the cleaning was completed, the next step was investigating the data to see what kind of insights could be gained. The first thing I noticed was in regard to the breeds that had been assigned to the images, the most common breed was a Golden Retriever. This doesn't entirely match up with the rankings from the American Kennel Club for 2017, but it is close. According to their 2017 rankings, the most popular breed that year was the Labrador Retriever, with the Golden Retriever being ranked 3rd. Considering the imprecise nature of machine learning, it is reasonable that some that may actually be Labs were designated as Goldens. (Most Popular Dog Breeds 2017)

Next I looked at the very scientific classification system that WeRateDogs uses. The four classifications of dogs are Doggo, Pupper, Puppo, and Floofer. I used a bar chart to show how the different classifications stacked up.

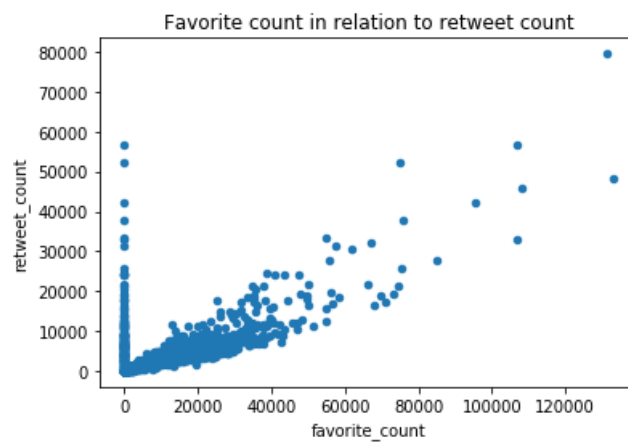
As we can see, the most common classification is Pupper. I did find it interesting that some dogs had more than one classification.

Next I wanted to see how the retweet and favorite counts compared.

This scatterplot shows that there is positive correlation between the retweet count and the favorite count. It also shows that tweets are frequently retweeted without being favorited.



Dogtypes.PNG



Tweetcount.PNG

With data we can come to a great many insights, and these are just a few regarding WeRateDogs. If you're not following them on Twitter you really should be. Everyone's feed could use more puppers.



(WeRateDogs 2022)

In []: