

DATA WRANGLE PROJECT

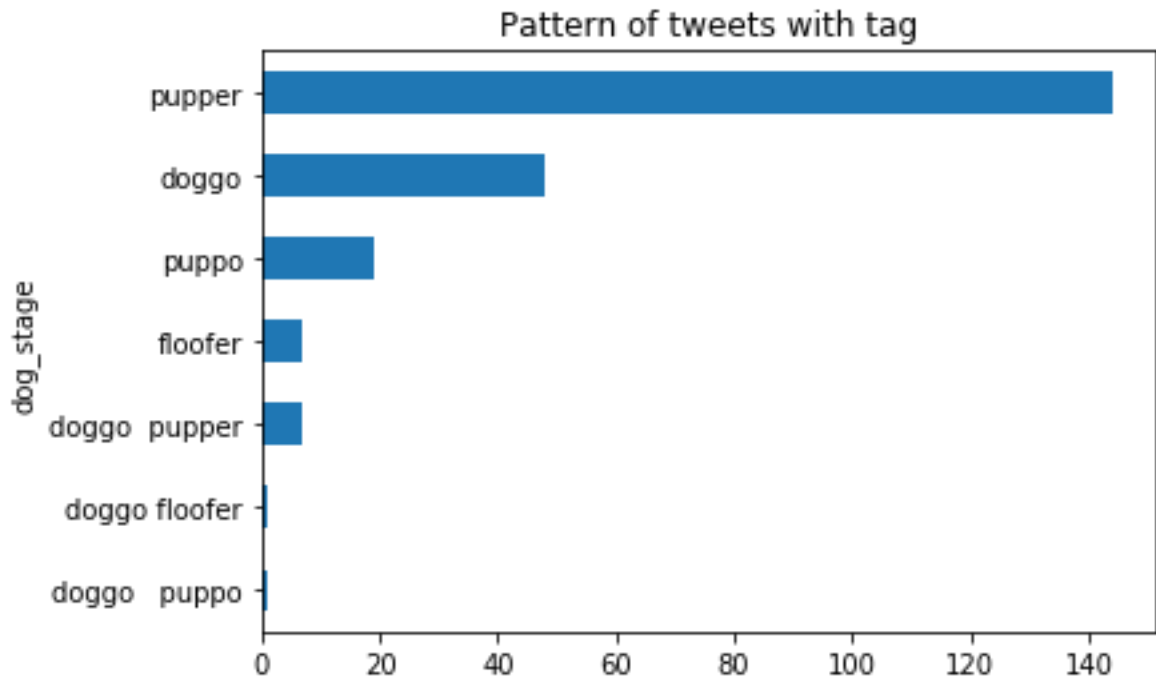
WERATEDOGS (Twitter data) – Analysis

Weratedogs is a twitter account that rates people's dogs with humorous comment about the dogs. It was started in 2015 by college student Matt Nelson, and has received international media coverage both for its popularity.



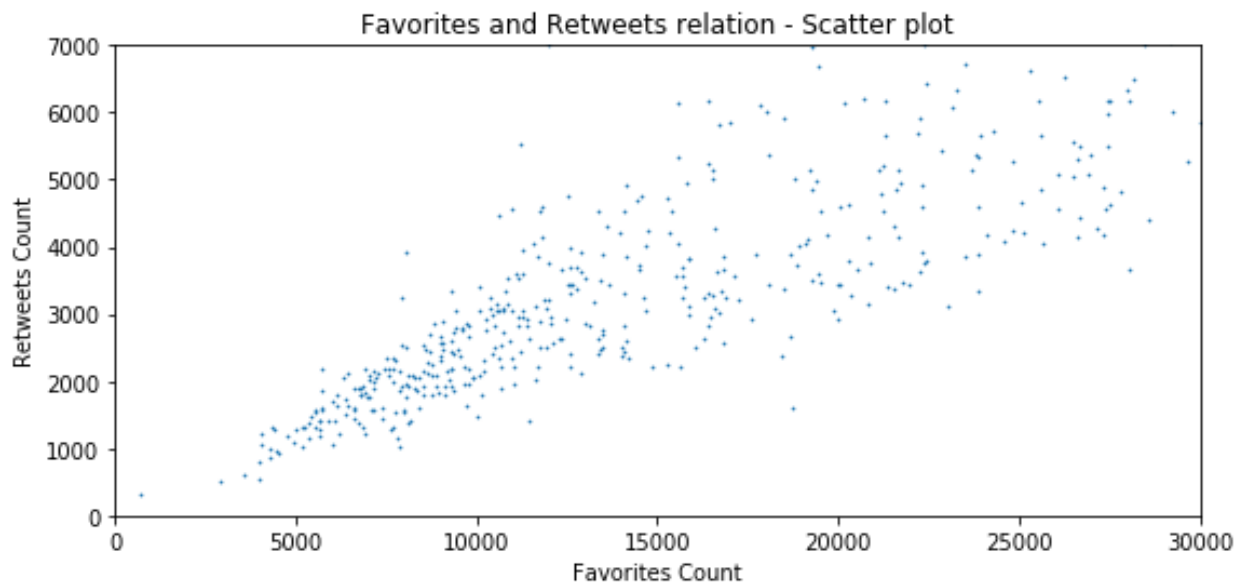
From this project, we aim to wrangle WeRateDogs Twitter data to create interesting and trustworthy analyses and visualizations. The Twitter archive is great, but it only contains very basic tweet information. Additional gathering, then assessing and cleaning was required.

1) Data from twitter archive to see dogs stages



Observation: "Pupper" stage dog has highest number of tweets with images in the account when compared to other stages. But based on this alone we cannot decide if this is the most loved stage.

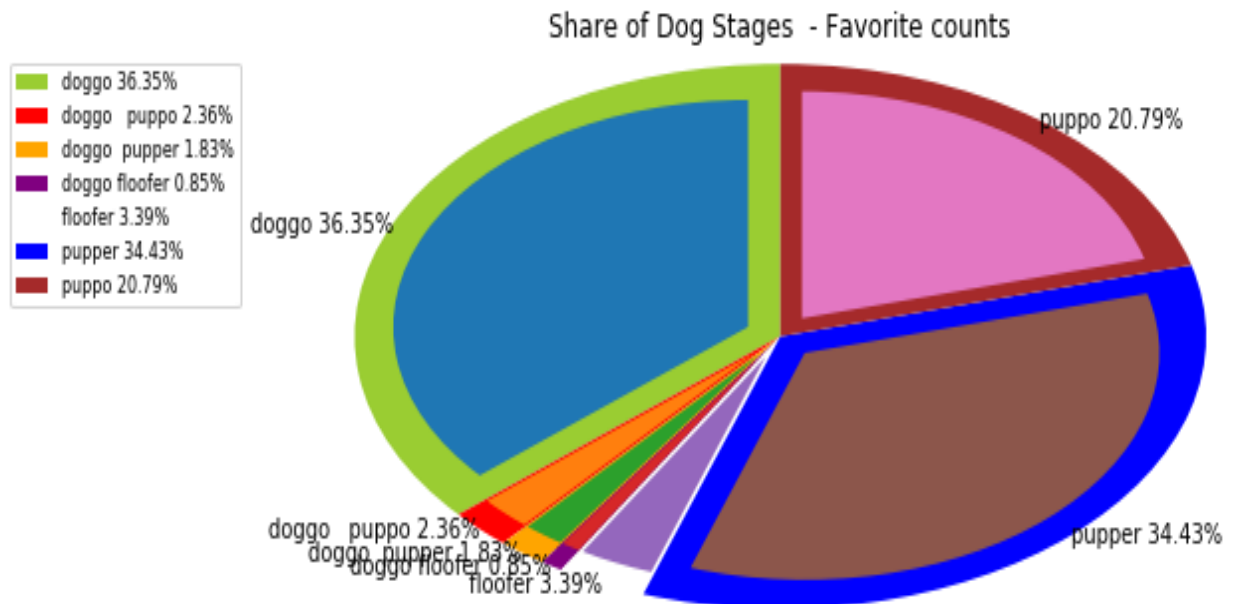
2) We now see the relation between retweets and favorites by deriving a scatter plot



We can observe that there is a positive correlation between the number of retweets and Favorites in the handle which means higher the retweets higher the favorites.

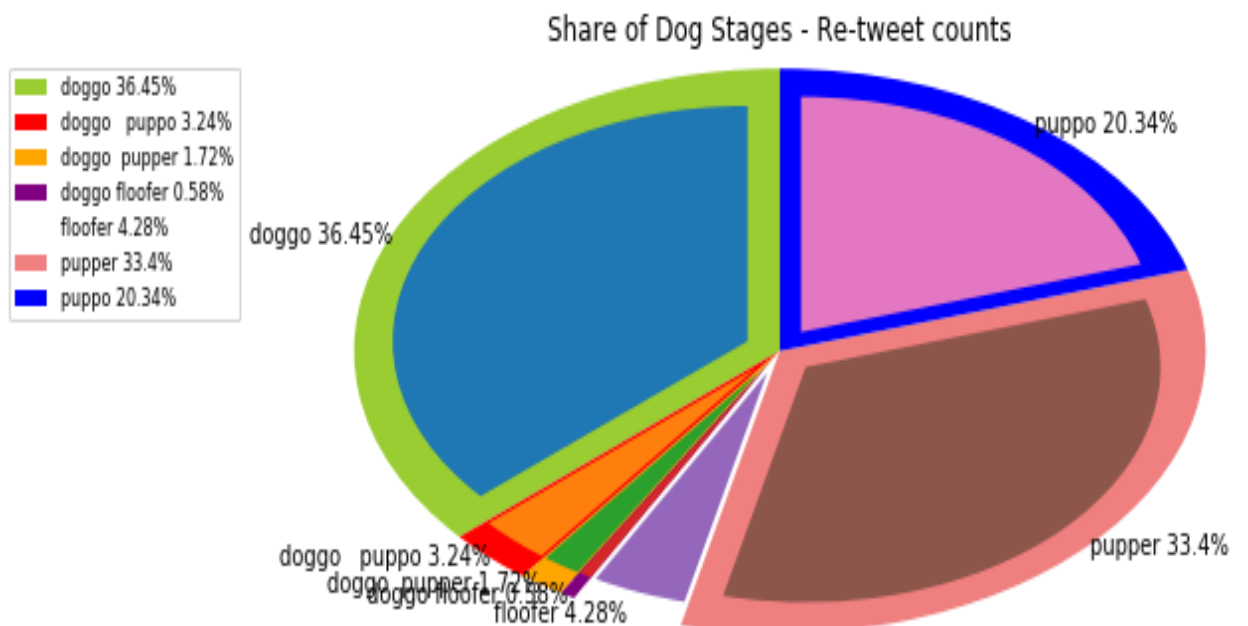
We will now use these two parameters to see which one is the better stage amongst the available ones

3) Check the Favorite count by dog stage



We observe "Doggo" stage is the highest in favorites
Continue further to check the retweet count as well

4) Check the retweet count by dog stage



Conclusion:

We can see that the Doggo stage is the highly re-tweeted, liked and favourited as well.

We also observe that # tweets from the archive file is having more records for "Puppo" (might be because the # records pulled from twitter api is very less (888))

We also observe that the "Golden Retriever" is the most tweeted dog breed amongst all the other breeds.

Finally, we can see that there is possitive corelation between the favourites and re-tweets which in turn can be depicted as if the number of re-tweets increase then there is possible chance of favourites and vice versa