
Visualization and Insight report for WeRateDogs

By Sarah Ali

Twitter account “WeRateDogs” provides ratings for dogs along side some interesting and hilarious remarks about them, these dogs are rated over 10, but the numerators are slightly over 10, so the ratings usually go like 11/10, 12/10, 13/10 and so on.

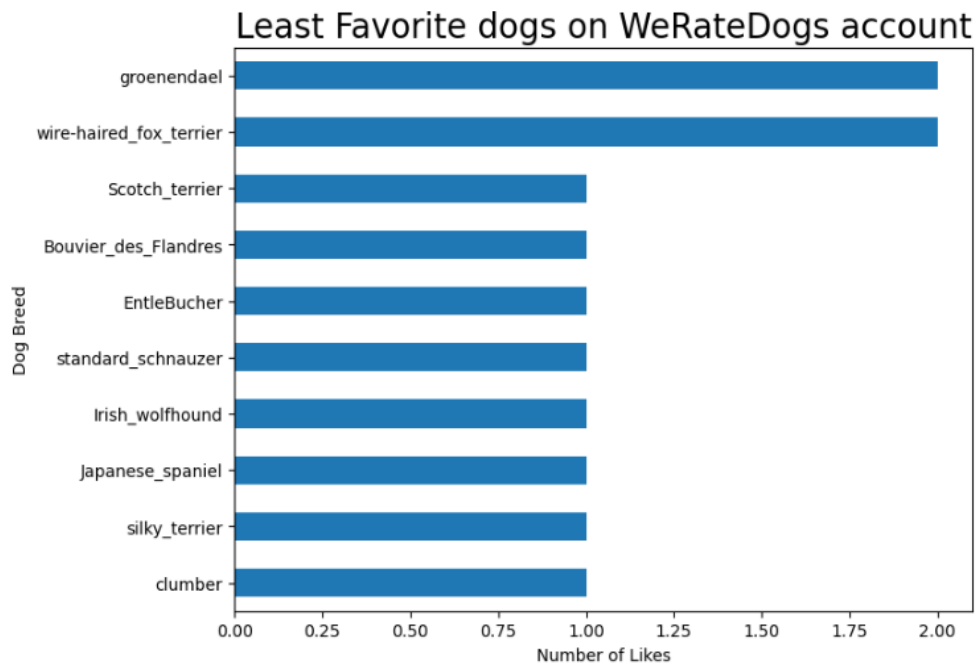
WeRateDogs downloaded their Twitter archive and sent it to Udacity via email exclusively to be used in this project. This archive contains basic tweet data (tweet ID, timestamp, text, etc.) for all 5000+ of their tweets as they stood on August 1, 2017.

In the dataset wrangling efforts were gathering, assessing, cleaning and finally combining all three datasets into one master document to be used in data analysis and visualization.

After a thorough wrangling session, here are few insights I want to share with you.

1. **What is the average rating amongst the dogs:** This was achieved by finding the mean values of all ratings per dog stage numerators since the denominators were all 10. This would result in:

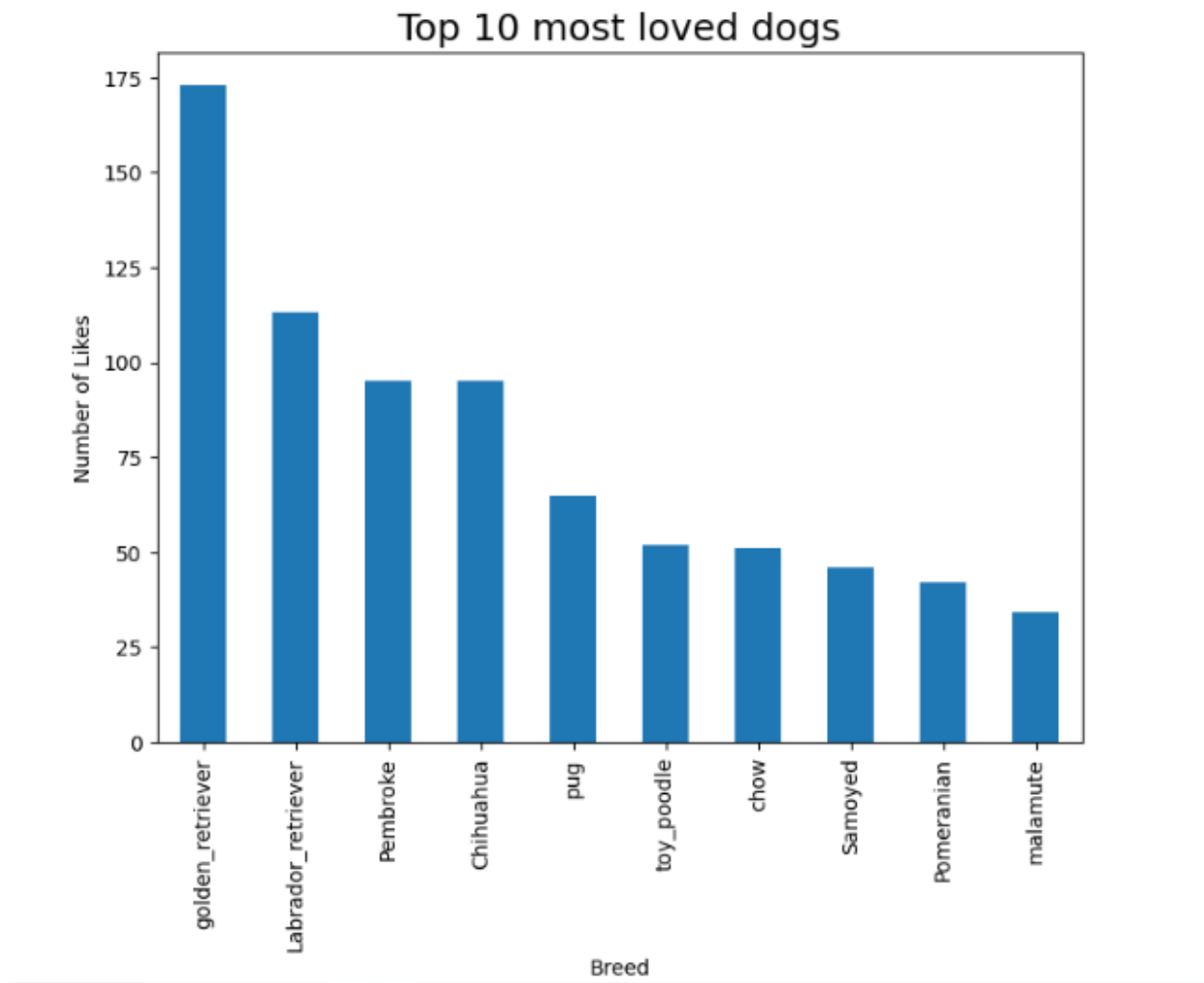
dog stage	Average rating over 10
None	10.803254
doggo	11.925926
doggo, floofer	11.000000
doggo, pupper	11.000000
doggo, puppo	13.000000
floofer	12.000000
pupper	10.849057
puppo	11.952381



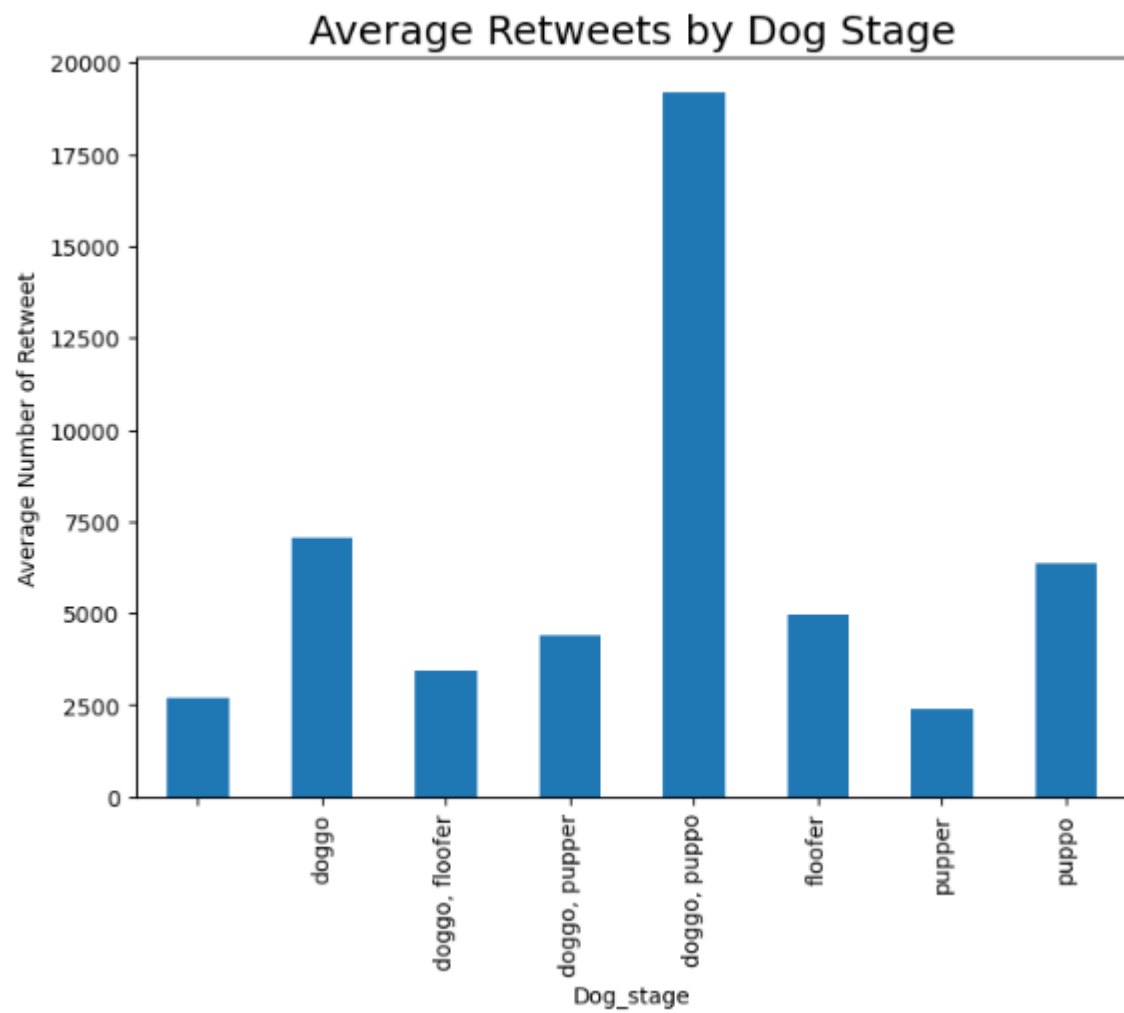
2.

The least likes dog breeds were 8, i.e. Clumber, silky terrier, Japanese spaniel, Irish wolfhound, standard schnauzer, Entlebucher, Bouvier des_Flandres and scotch terrier, who would have thought.

3. **The most liked dog breed:** was the golden retriever, my personal favorite.

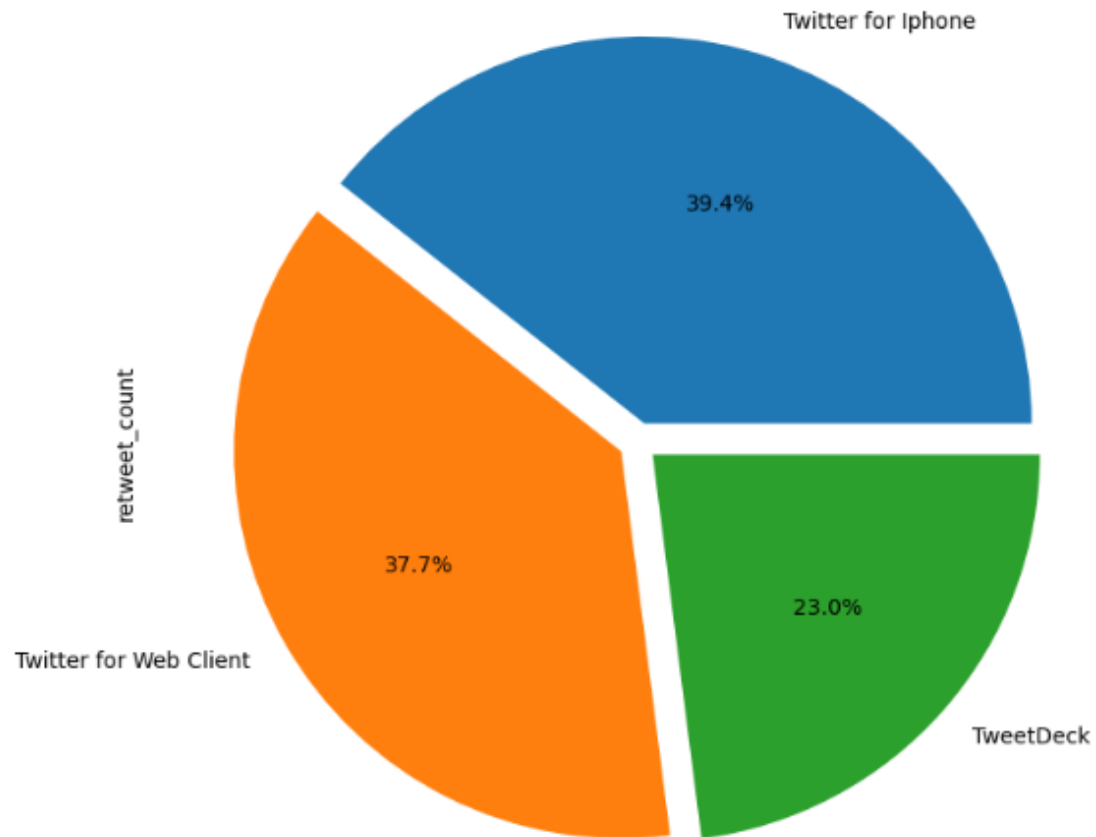


4. **The average number of retweets for each dog stage:** Dogs in stages doggo, puppo or both of them doggo, puppo seems to have the highest retweets so far. The counts for the dogs without dog stage is rather high to come to a conclusion whether or not the dog stage of the dogs' influences retweets.



5. The source of these picture has any influence on whether the audience likes the dogs

The source of tweet and how many times it was retweeted



6. **Is there a relationship between the likes and retweet:** Most favorited tweets were retweeted and vice versa. There is a strong correlation between the likes and retweets. The regression coefficient for this relationship is strong ($r = 0.788$).

Correlation between Retweets and Likes

