Analysis Report: WeRateDogs Twitter Data

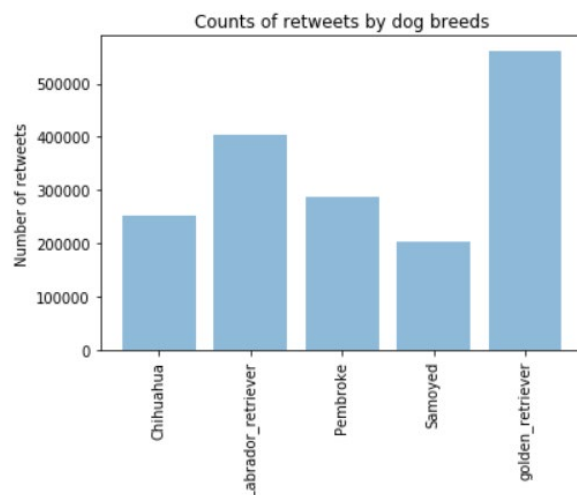**By: Quang Luong**

Date: 12/9/2018

Analysis:

The following analysis is from the Twitter data, called WeRateDogs. This analysis is mainly concentrating on the ratings, retweet count, and favorites count by dog breeds.

1) <u>Count of Retweets by dog breeds</u>:
   a. Since there were over 100 dog breeds, I filtered it to where only those dog breeds that have at least 200,000 retweet counts.
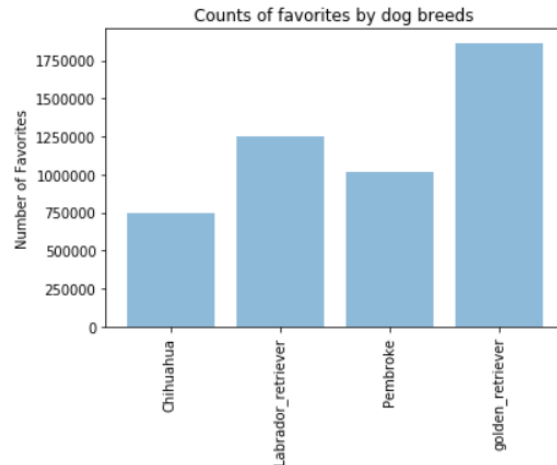   b.

Counts of retweets by dog breeds

   c. From the above we could see the breed with the most retweets are as follows:
      i. Golden Retriever
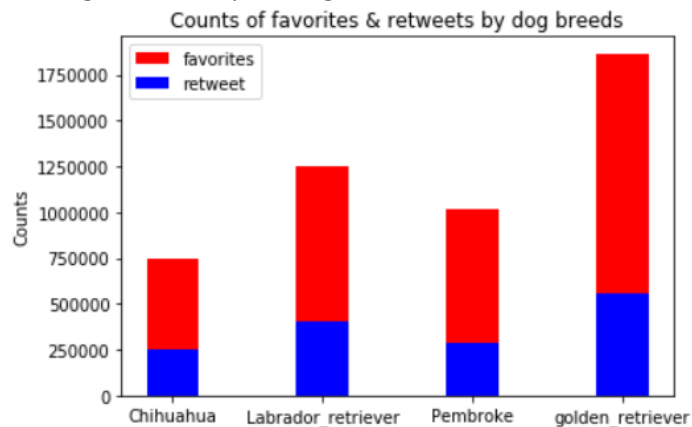   ii. Labrador Retriever
      iii. Pembroke

2) <u>Count of Favorites by dog breeds</u>:
   a. Again, since there were so many dog breeds, I narrowed the list to those that have at least 700,000 counts.
   b.

Counts of favorites by dog breeds

c. From the above we could see the breed with the most Favorites are as follows:
  i. Golden Retriever
  ii. Labrador Retriever
  iii. Pembroke
d. Not too surprising that since a mark for a favorite dog breed also resulted in an action for retweet.
e. Looking at the two plots together we can see the same order:

Counts of favorites & retweets by dog breeds

3) Average Ratings by dog breeds:
  a. For the average ratings, again, since there were too many dog breeds, I took just the top 10 highest average rating:

	dog_breed	rating_numerator
53	Saluki	12.500000
77	briard	12.333333
62	Tibetan_mastiff	12.250000
103	silky_terrier	12.000000
106	standard_schnauzer	12.000000
30	Irish_setter	11.750000
25	Gordon_setter	11.750000
54	Samoyed	11.731707
21	Eskimo_dog	11.600000
80	chow	11.522727

b. What I find interesting here is that the Golden Retriever, which had the highest retweets and the highest favorite counts, did not get the highest average rating. The highest average rating is the Saluki.

    i. In fact, with the exception of the Samoyed, none of the other dog breeds from the top retweets or favorite is even in this top 10 list.