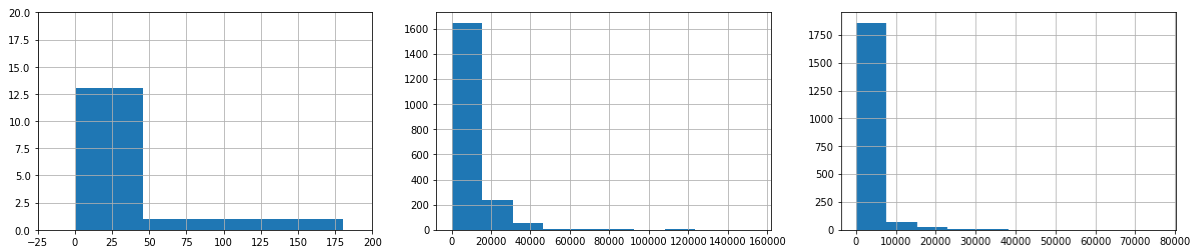# WeRateDogs Data Wrangling
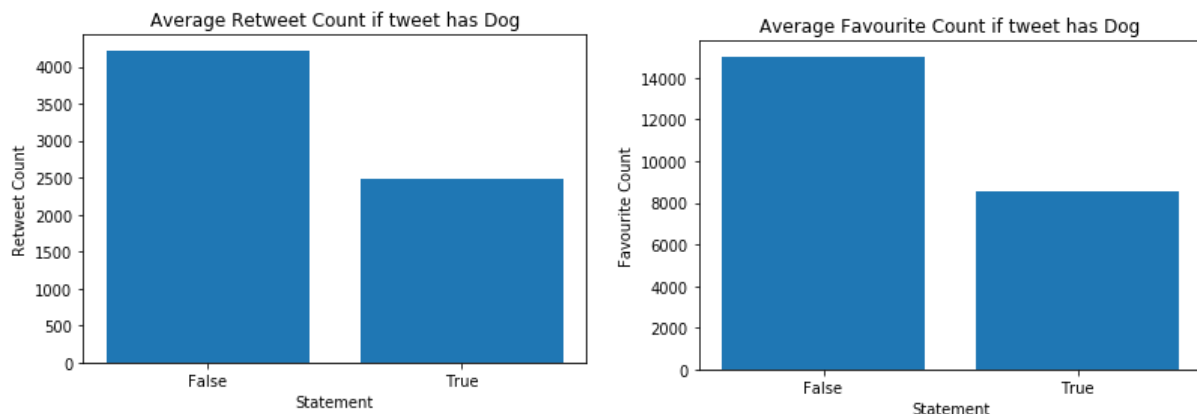
Insights and Visualisations:

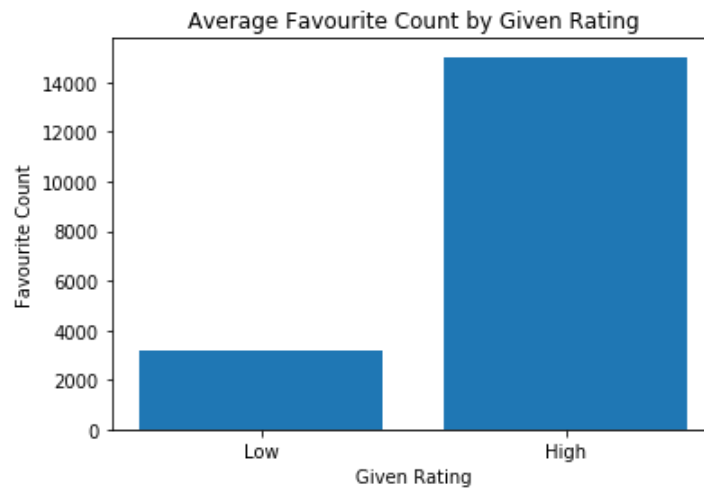**Numeric values skewness: (Rating - Favourite Counts - Retweet Counts)**

- All three variables are positively skewed, most of the data are below the average point.

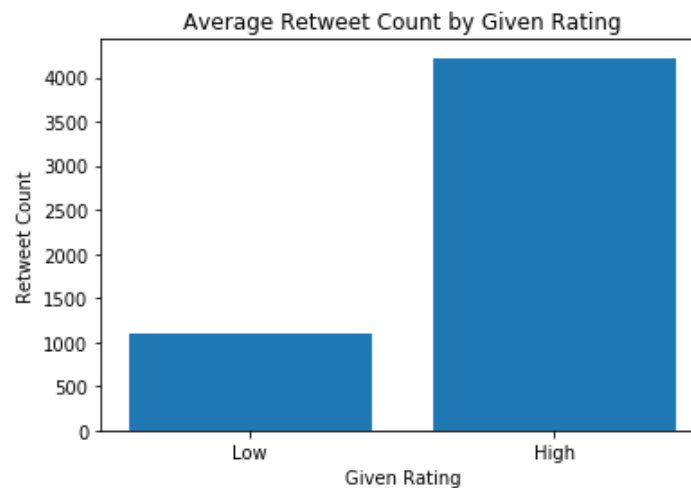**Does tweet content affect retweet and favourite counts ?**

- Despite that WeRateDogs is actually a dog based page, the numbers using the prediction algorithm suggests that people interact more with tweets that do not have dogs as a part of them.
- This was calculated by separating data into 2 sets, one where prediction is a dog, the other is not, then using the average count of every set.

Do tweets with higher ratings receive higher favourite counts ?
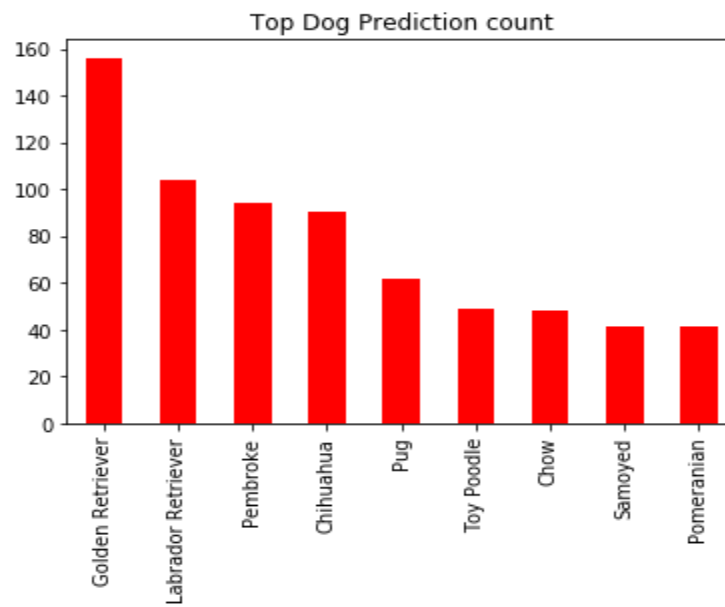
Average Favourite Count by Given Rating

- This was calculated using the median of ratings then dividing the data onto 2 separate, low and high, sets, as shown, the rating number actually affects the Tweet's Favourite Counts tremendously.

Do tweets with higher ratings receive higher favourite counts ?

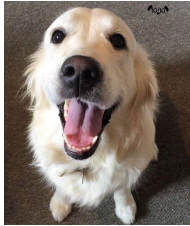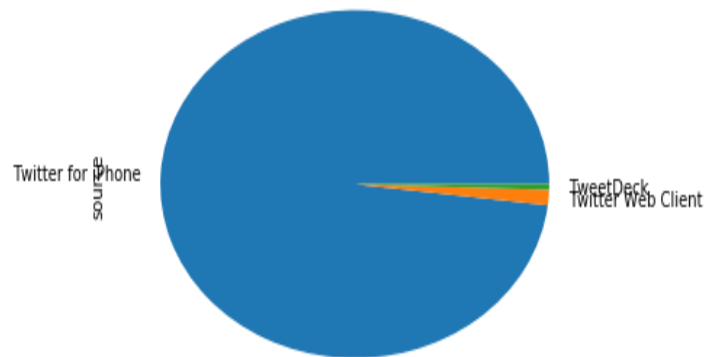Average Retweet Count by Given Rating

- This was calculated using the median of ratings then dividing the data onto 2 separate, low and high, sets, as shown, the rating number actually affects the Tweet's Retweets Counts tremendously.

What are the most frequent Dog breeds from predictions ?

- The data shows these exact numbers of predictions for the top 9 breeds found in dataset 156 Golden Retriever, 104 Labrador Retriever, 94 Pembroke, 90 Chihuahua, 62 Pug, 49 Toy Poodle, 48 Chow, 41 Samoyed, 41 Pomeranian.

A chart for different source of tweets

- It seems from this chart that the admins mostly use Iphone while creating new tweets.