Act Report - WeRateDogs Twitter Analysis

Date: 2/8/2019   By: Alex Cox

Insight 1: Quantity of favorites per dog name

The total quantity of favorites per dog name show there are a few dog names appearing frequently in tweets from the WeRateDogs account that may be skewing the data due to high frequency of use. Average retweets does not appear to be skewed by the frequency of use per dog name in tweets. None of the top 10 most often used dog names in tweets appears in the top 10 dog names by average favorite count,

Insight 2: Quantity of retweets per rating

The most frequently used ratings appears to be base 10, with 1-14/10 comprising 1952 of the 1975 or 98% of ratings used in the tweets. Predictably, total quantity of retweets follows the same pattern of the frequency of use of ratings: the top 10 total quantity of retweets per rating is comprised of base 10 ratings from 13-6 / 10, which I would expect to see purely based on the frequency the ratings are used. Also predictably, when we look at the average of retweets per rating, the top 10 is comprised of 70% of ratings with low frequency of use in tweets, showing us that their raw number of retweets is relatively high in comparison to the average of the other more often used ratings. The top 10 also shows some "vanity" ratings, e.g. 9/11, 7/11, 420/10, that may have caused a skewed uptick in the quantity of retweets despite their low frequency of use. The notable exception is 14-12/10, given their extremely high frequency of use (they are used ~64% of the time), their average retweets are higher.
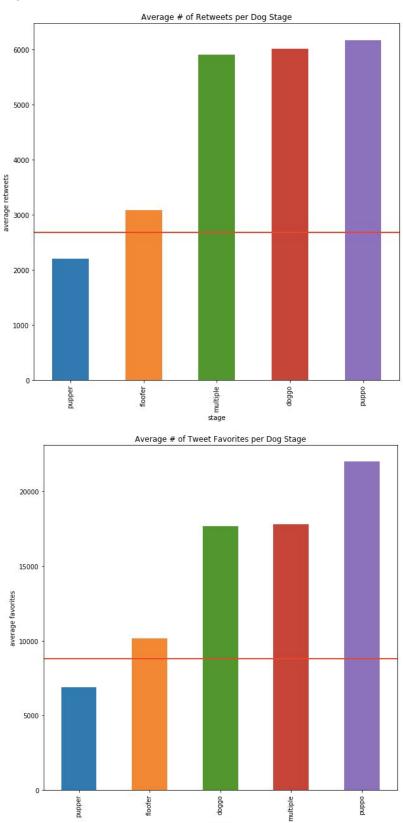
Insight 3: Quantity of favorites per rating

Favorites per rating tracks nearly identically to retweets, with 14-6/10 occupying the top 10 total favorites and 'vanity' and alternative ratings occupying most of the top 10 average quantity of favorites. The average favorites of the dataset is ~8806, the top 9, or 27% of the ratings average number exceed the average.

Visuals: Quantity of retweets/favorites per dog stage

I evaluated the retweets and favorites per dog 'stage' category and found on average the tweets indicating a dog is of a 'puppo' stage on average narrowly has the most retweets and by far the most favorites. "Doggo' and tweets with multiple stages are 2nd and 3rd most retweeted/favorited tweets on average. The WeRateDogs account uses many iterations of the terms, as part of cleaning I reset all like values to the normal stages, for example 'puppoccino', 'pupporting' were set to 'puppo'. Some tweets used multiple terms, which I set to 'multiple'.

Looking at the raw values, 'puppo' seems to have its average dragged down by the sheer quantity of times it appears in tweets since it is the lowest, while puppo has fewer mentions, but much higher average retweets and favorites.

Average # of Retweets per Dog Stage

Average # of Tweet Favorites per Dog Stage