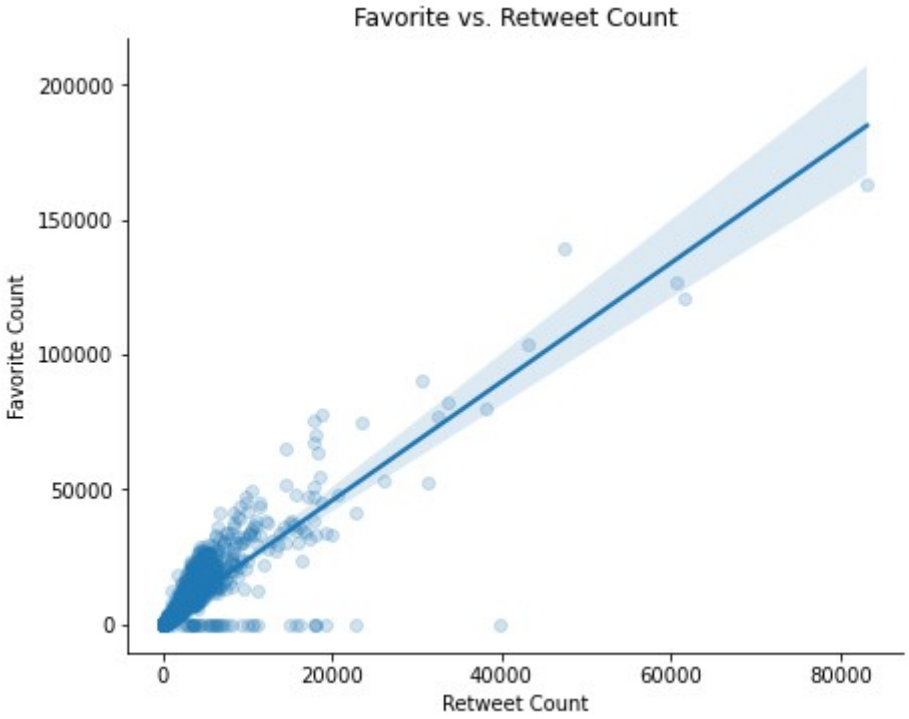
Project 2 : WRANGLE and ANALYZE Data

The dataset that you will be wrangling (and analyzing and visualizing) is the tweet archive of Twitter user dog_rates, also known as WeRateDogs WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. Why? Because "they're good dogs Brent." WeRateDogs has over 4 million followers and has received international media coverage.

Project Steps Overview

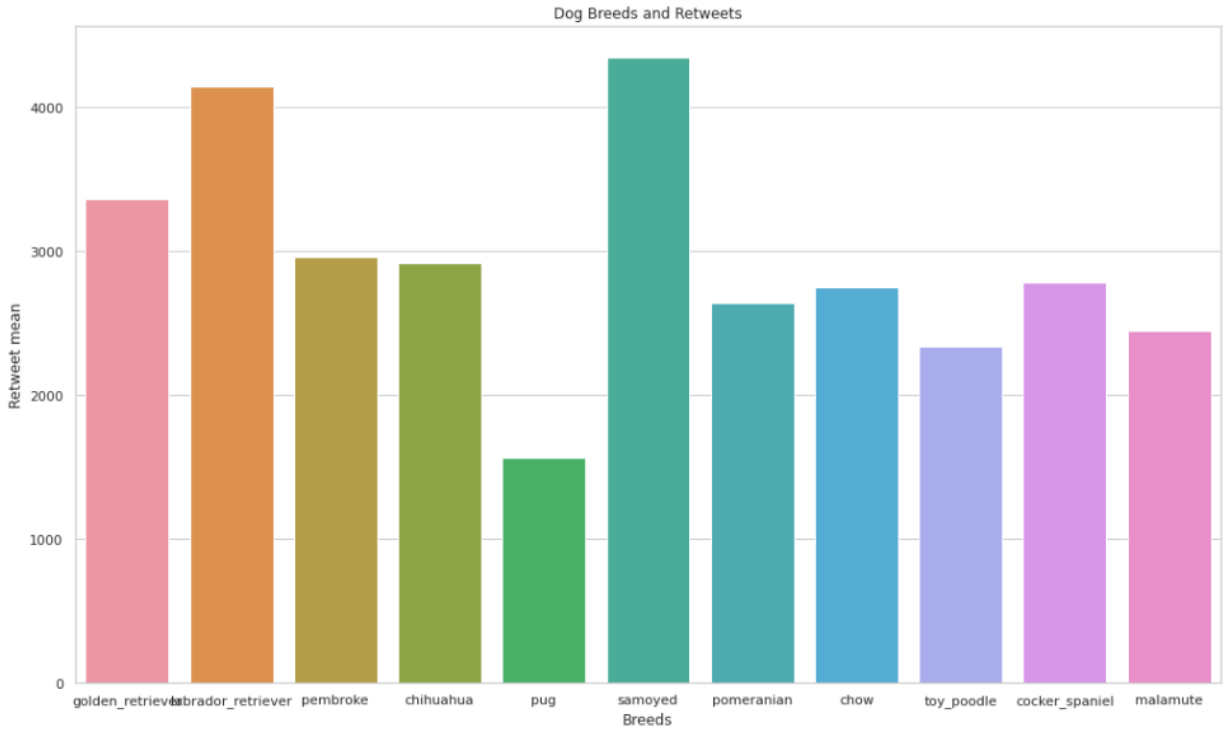
Analyzing, and visualizing data

What Questions Are We Trying To Answer?



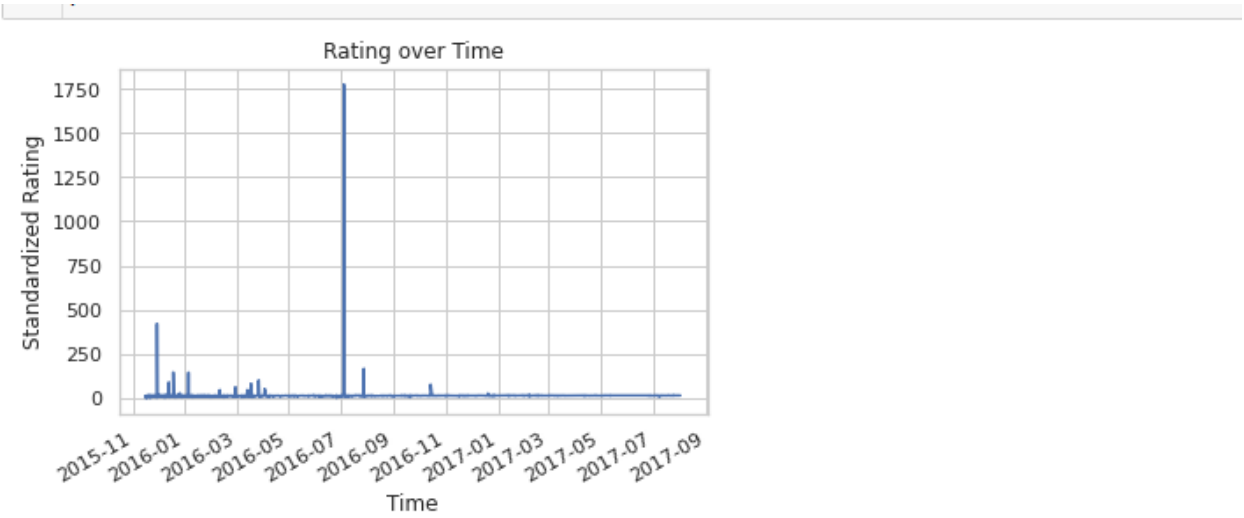
From the above visualization the Favorite count and retweet counts are highly positively correlated. For about every 4 favorites there is 1 retweet. The majority of the data falls below 40000 favorites and 10000 retweets.

2. Which are the more popular doggo breeds?



The samoyed and the golden_retriever have the highest retweeted post compared to with the other breeds

3. Are the ratings Standadized ?



the rating in period of 2016-07 is highly un-standardize.