

# Analysis and Visualization Report

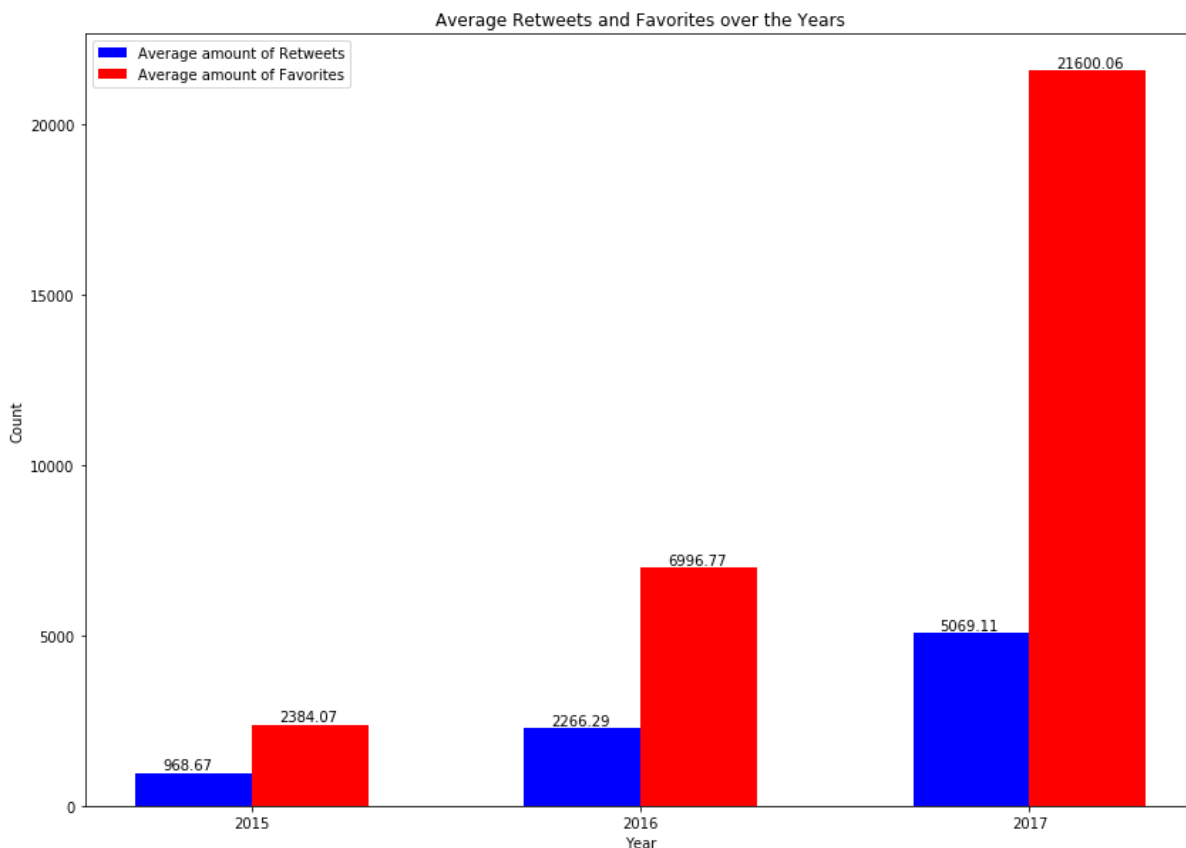
## Supporting the Data Wrangling and Analysis of the Twitter account WeRateDogs.

This report will be complementary to the Data Wrangling that was performed for the Twitter account WeRateDogs. All the Data Wrangling, Analysis and Visualization was done in the notebook `wrangle_act` that is in this workspace for further understanding. Data Wrangling involves 3 steps, first is to Gather, then Assess and finally Clean which can be found in `wrangle_act`. Once we have all of our clean data, we are able to perform analysis and gather some insight into our investigation. During the process of working through and getting to know our data, I've decided to try and answer 4 different questions.

- How do favourites and Retweets Change over the Years?
- Distribution of Tweet Sources.
- Are Lower Ratings Given to Non-Dogs only?
- Which Dog Breed was the most Predicted

### How do Favourites and Retweets Change over the Years?

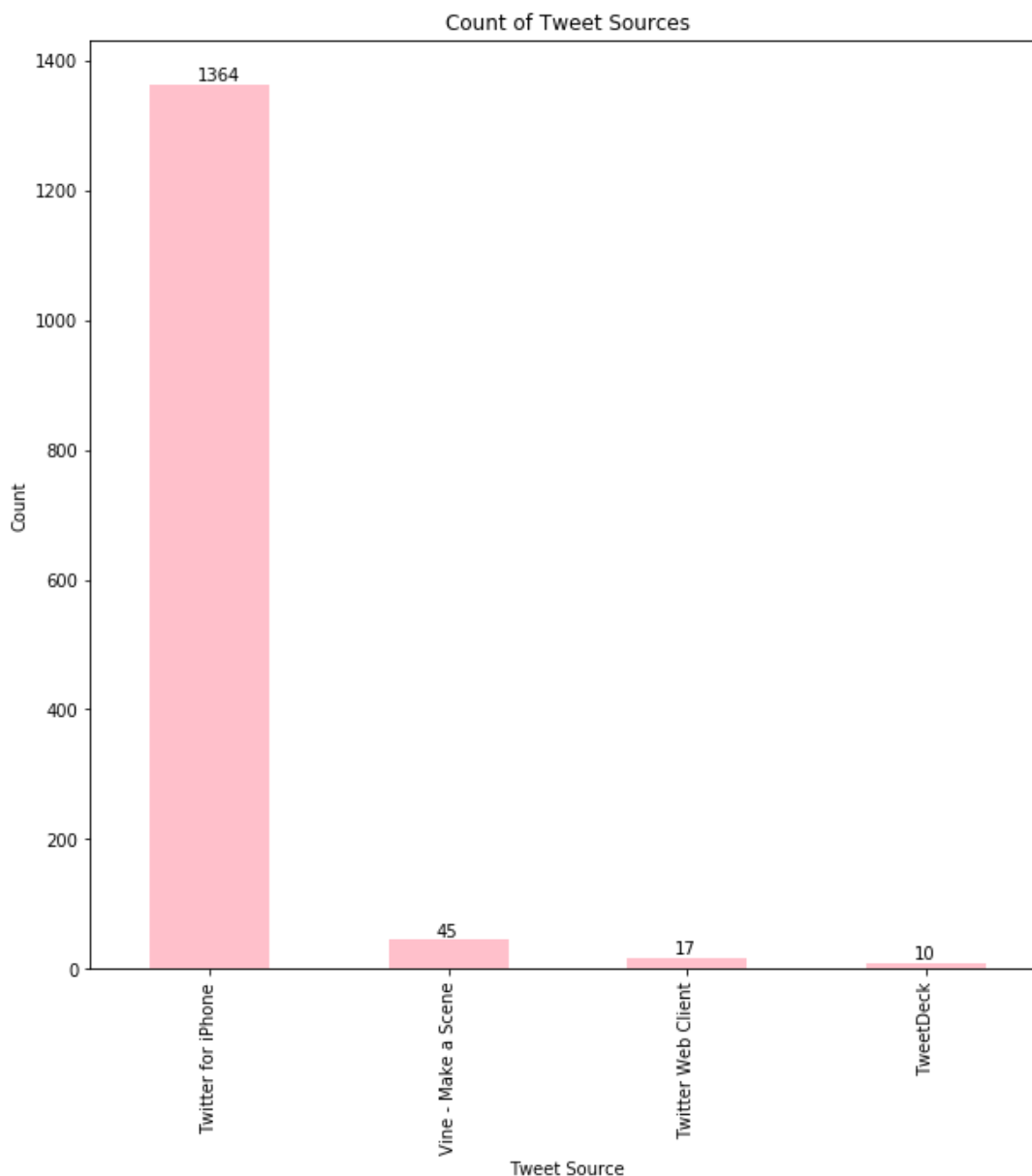
When you favourite a tweet, it is synonymous to a "like" and will be saved into a list of your favourited tweets on your profile. When you retweet a tweet, it is being posted on your own timeline where your followers are also able to see the tweet from the original tweet author. I wanted to see how these two changed over the years and knowing that this account has 4 million Twitter followers, I thought it would be interesting to see where most of their activity comes from.



As you see from the bar graph above, 2015 has the lowest amount of favourites and retweets and the year 2017 has the most. The amount of favourites has increased drastically in 2 short years, and retweets also steadily increasing. As the account gains more followers, the amount of activity on their account would also increase.

## Distribution of Tweet Source

From our data, we were able to extract the 4 different types of sources WeRateDogs used to tweet. Twitter for iPhone- which is most likely an app, Vine, Twitter Web Client and TweetDeck. Vine is an app used for short videos at 6 seconds long. Twitter Web Client is essentially using Twitter in a web browser and TweetDeck is a dashboard app that is able to manage your Twitter. You can see from the chart below that Twitter for iPhone has taking nearly 95% of all the sources. It's clear from the results that Twitter for iPhone is the most popular and most frequently used to tweet for WeRateDogs.

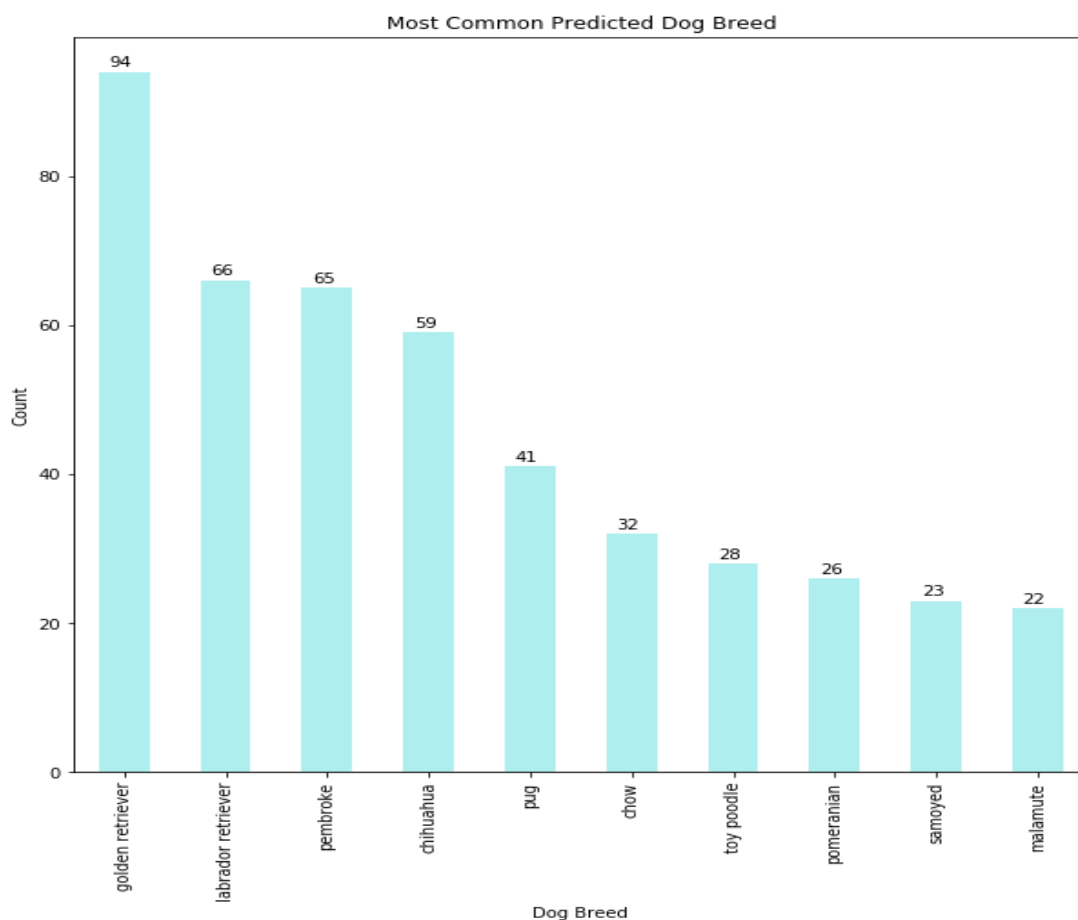


## Are Lower Ratings Given to Non-Dogs Only?

Seeing as this account is a humorous Twitter account that has gained it's 4 million followers through funny and adorable captions with dogs, it would seem counterintuitive to have low ratings given to any of the dogs in the tweets. When I was assessing the data, I noticed some tweets that had low ratings, 2/10, 3/10 etc. which I thought was odd seeing the nature of this account. Upon some quick investigation, I saw that the tweets that had low ratings contained images of non-dogs. These included birds, lizards, goats, the Target dog statue, which were all obviously not images of dogs for a dog rating account. So instead of going through all the tweets with low ratings, I decided to investigate programmatically. Using the image predictions on these tweets, I calculated that 67% of the tweets that had low ratings were all not predicted to be images of dogs. This helped me with some insight and affirmed my question that lower ratings were predominantly given to non-dogs.

## Which Dog Breed was the Most Predicted?

Since we had a model that predicted all the dog breeds in these tweets, I thought it would be interesting to see which dog breeds were the most predicted/ which dogs were most frequently shown on this account. As you can see below, the golden retriever and the labrador retriever were the two most popular/frequently predicted dog breed. These results make sense as labrador retrievers are the most common dog breed, followed by german sheppards, and then golden retrievers. It would be interesting to see how accurate our model was, seeing as golden and labrador retrievers are very similar.



## Conclusion

In conclusion, after data wrangling, to get more in depth and understand our data more, we were able to analyze and produce some visualizations to back up our insights. We learned about the Data Wrangling process, which you can follow in `wrangle_act` notebook and also see the visuals that were shown here, produced over there.