**act_report.pdf**

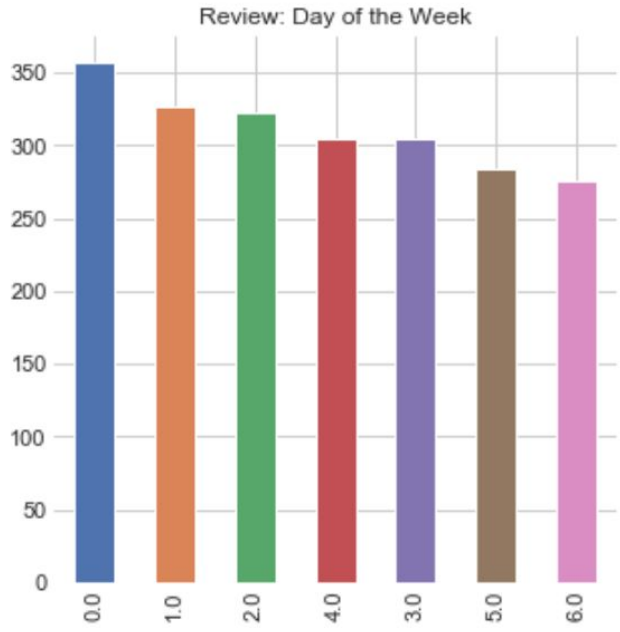
## INTRODUCTION

In this project, we gathered, reviewed, and cleaned data related to the WeRateDogs Twitter archive, @dog_rates, also known as WeRateDogs. The final data includes the WeRateDogs Twitter archive, image predictions from a machine learning model, and tweet data gathered using the Twitter API wrapper, tweepy.

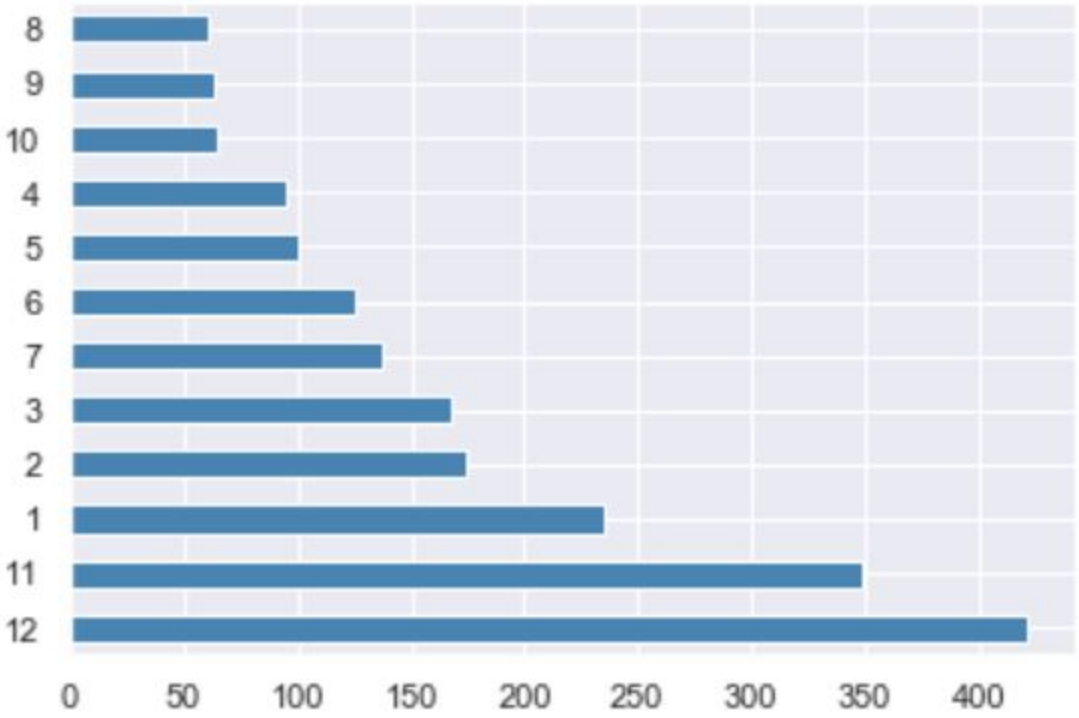
WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. WeRateDogs has over 4 million followers and has received international media coverage.

## FINDINGS

1. Once the data was gathered, cleaned, and merged, we ended up with 2356 entries, with 9 features.
2. The most popular dog name was Charlie, followed by Lucy, Oliver and Cooper.
3. Monday was the most popular day of the week for posting



4. December is the month with the most tweets to WeRateDogs. August has the least amount of tweets to WeRateDogs.

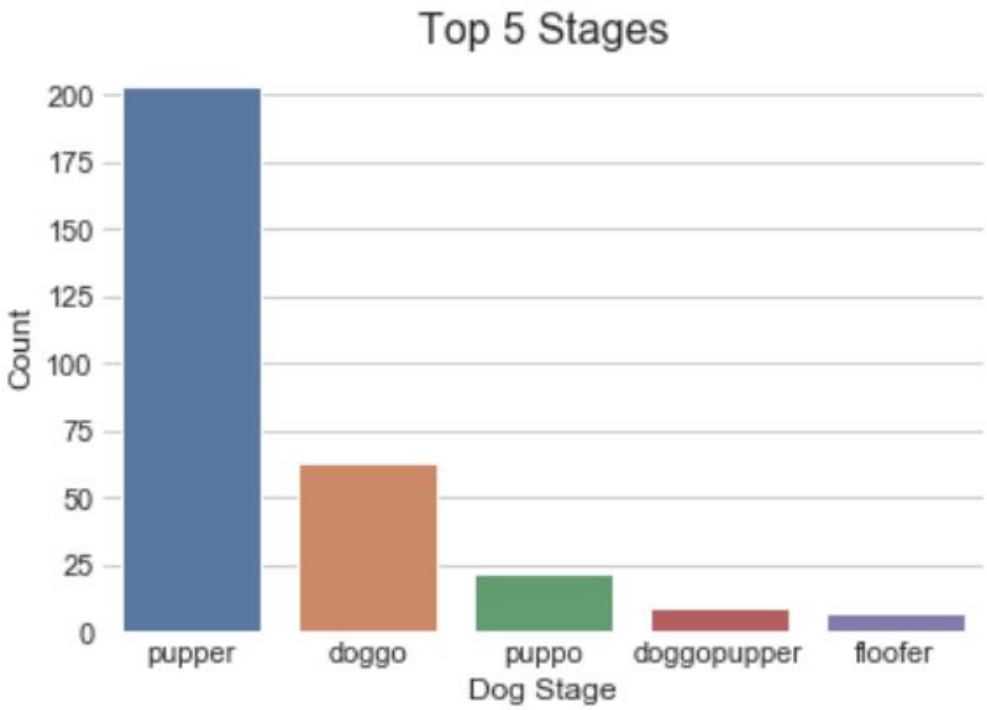



5. A majority of posts to the WeRateDogs account were from iphone users using the iphone Twitter app.

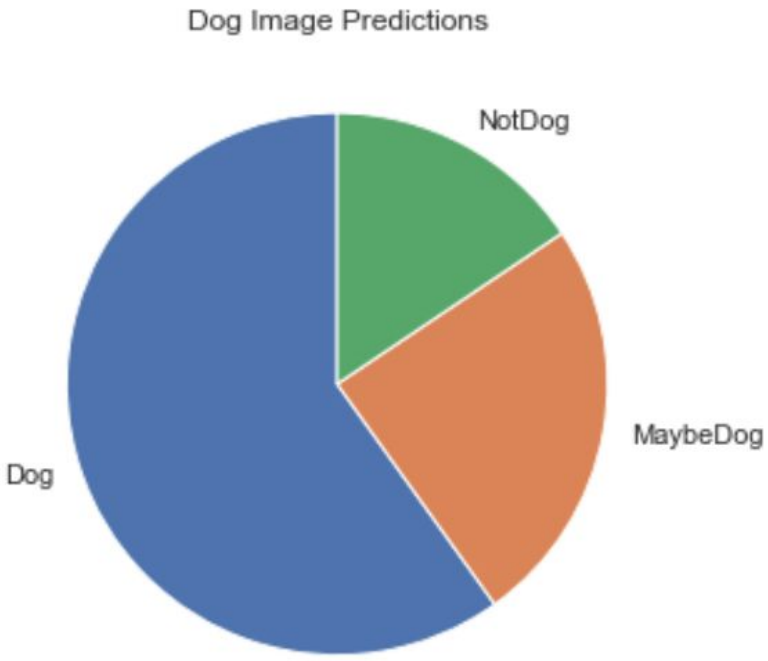
```
data.source.value_counts()
```

```
Twitter iPhone    1955
Twitter Web         28
TweetDeck           11
Name: source, dtype: int64
```

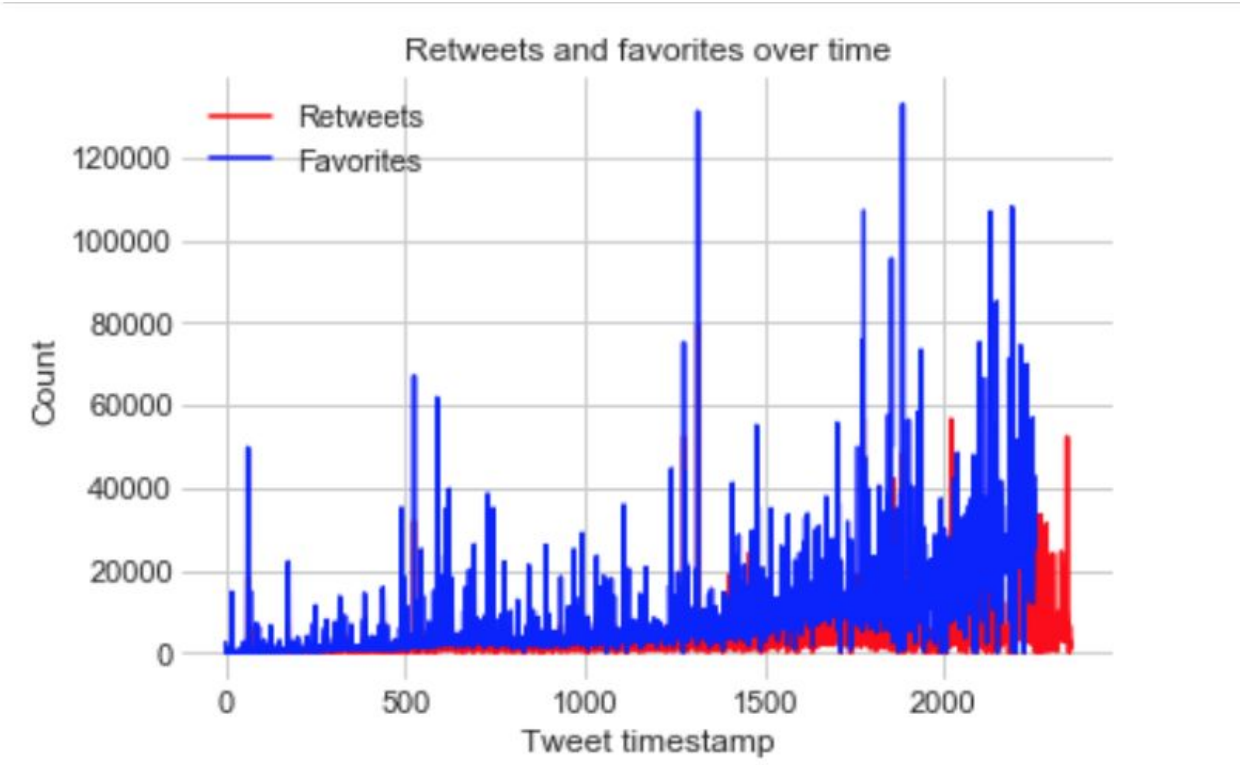
6. The majority of the dogs in the set were listed as ‘pupper’, a small, usually younger dog. Doggo followed this category, to take second place. The majority of the dogs, did not have a dog stage listed, as evidenced by the larger null blue column.



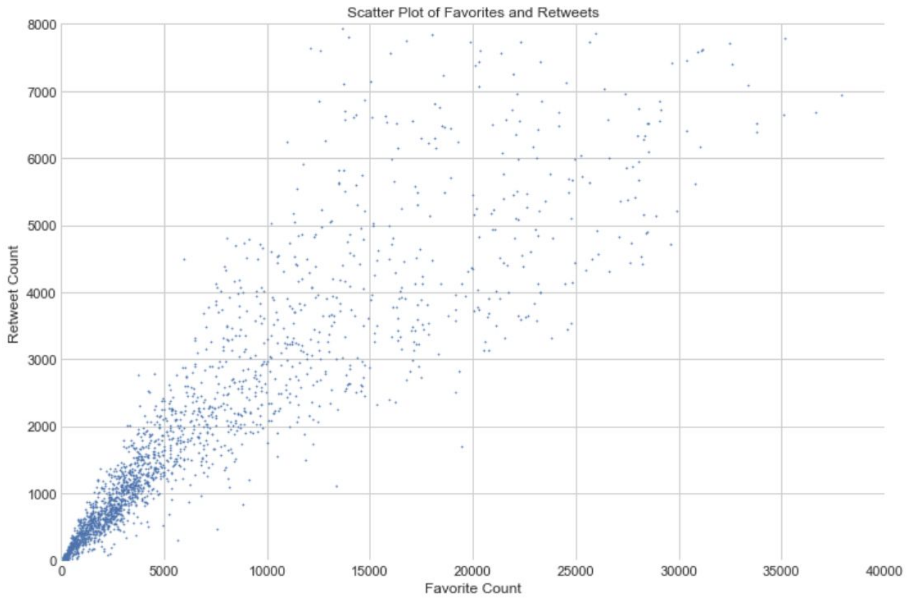

7. A large portion of the image predictions were dogs...but not all.



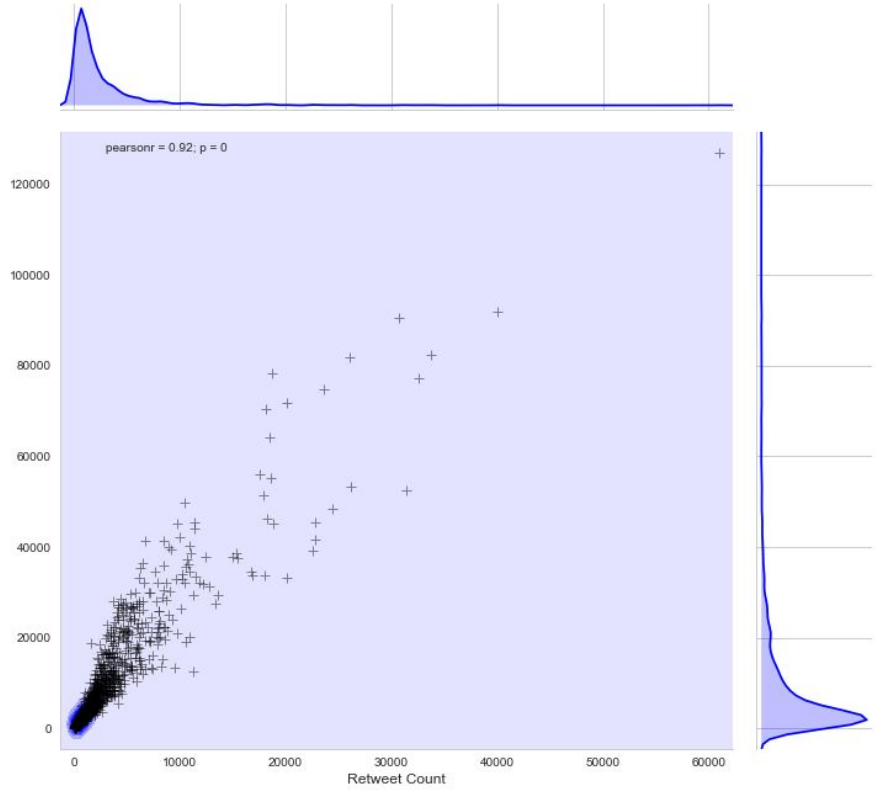

8. Entries generally had higher favorite counts than retweet counts



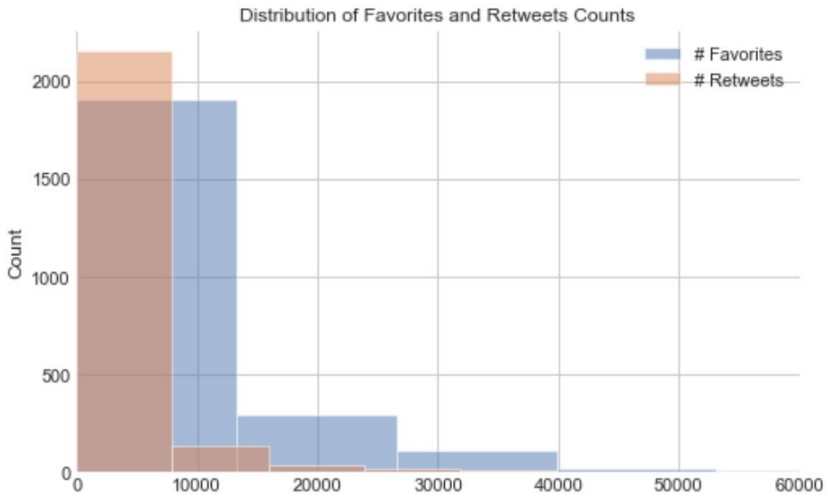

9. We see a positive correlation between retweet and favorite counts

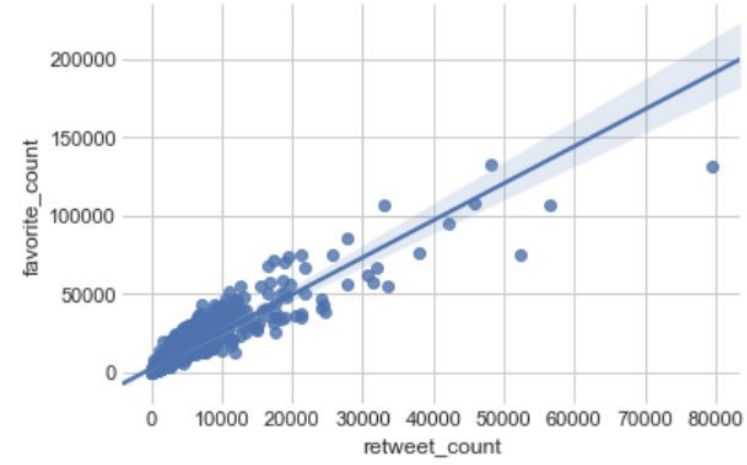



correlation review: retweet and favorite counts




10. Retweet and Favorite Counts follow a similar pattern


Distribution of Favorites and Retweets Counts
# Favorites
# Retweets
Count
2000
1500
1000
500
0
0
10000
20000
30000
40000
50000
60000


favorite_count
200000
150000
100000
50000
0
0
10000
20000
30000
40000
50000
60000
70000
80000
retweet_count