Analysis Report for WeRateDogs Dataset

By Hindhuja Gutha

In previous part of this project WeRateDogs twitter data is gathered from various sources and a clean data set is prepared which is analyzed . Below are some of the insights and visualizations

Insights :

1. **Most favorite tweet**

- It’s picture of a doggo realizing it can stand in a pool
- Above dog picture got highest number of favorite count 164884
- Algorithm predicted the type of dod as Labrador_retriever
- It’s rating is 13/10
- Name is not mentioned in the tweet

2. Most retweeted dog

- Most favorite tweet also got most number of retweets 84502

3. Least favorite dog

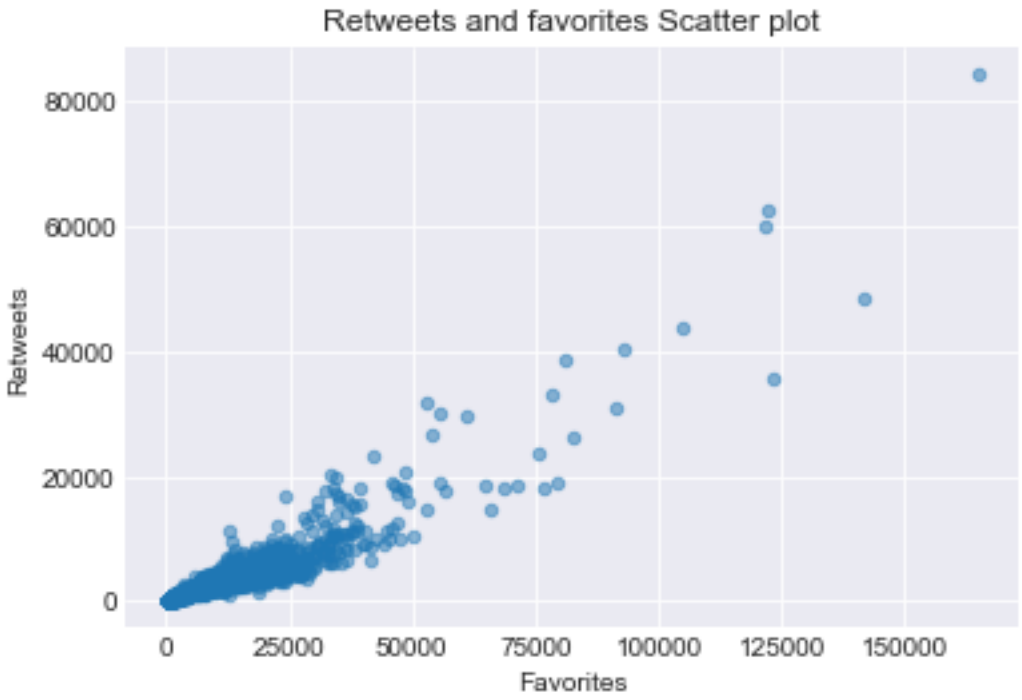
- favorite count for above dog is 80
- It’s an Adobe Setter
- machine learning algorithm predicted it as a English_setter
- Rating for this dog is 11/10
- Name is not mentioned in the tweet

3. Least retweeted dog

- Least favorite dog is also least retweeted dog with just 12 retweets

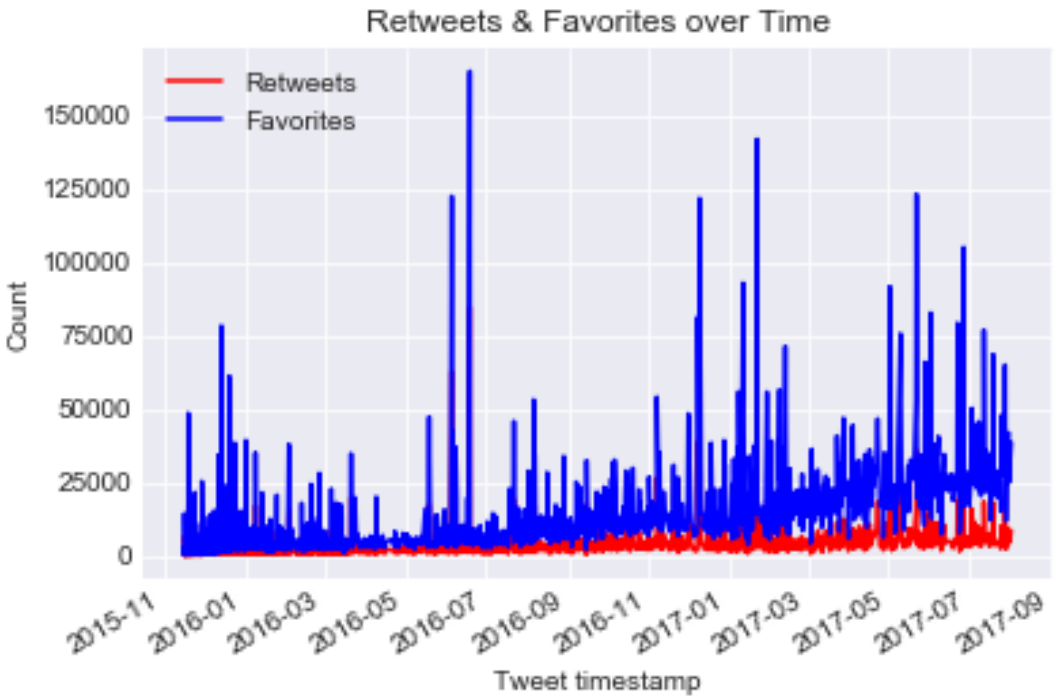
Visualizations :

1. Retweets and Favorites



- There seems to be a positive association between retweet count and favorite count
- Only few tweets are present with more than 50,000 favorites and 20,000 Retweets
- There are some outliers in this scatterplot
- As number of favorites and retweets increased count decreased for the tweets

2. Plot for Retweets and favorites Timeline :



- Above plot is drawn using retweet_count , favorite_count from data gathered using twitter API.
- In this plot we can see that there are some high spike in retweets and favorites
- Favorites are always greater than retweets throughout the dataset
- There is a overall increase in favorite count as well as retweet count