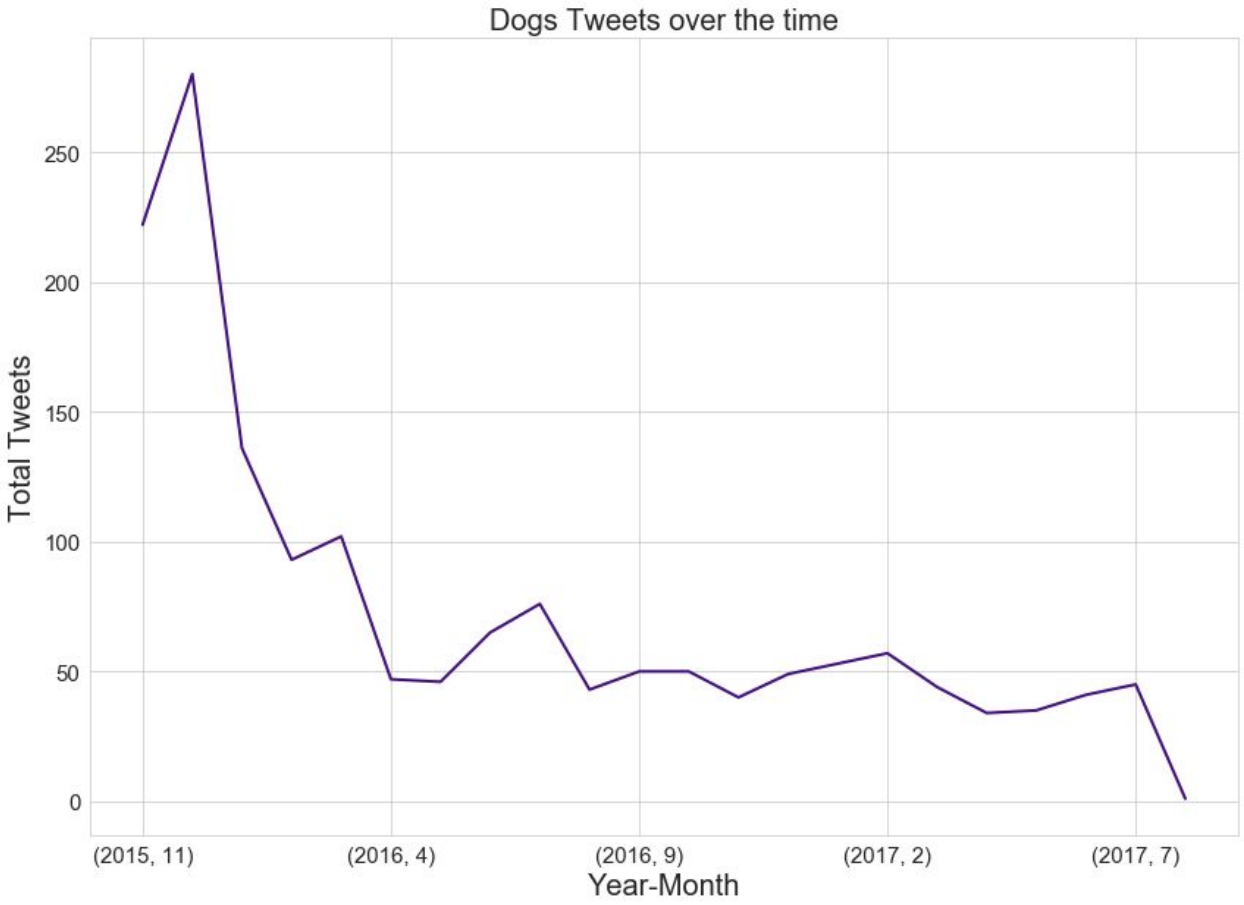
Data Insight and Visualization

I have analyzed the data with the help of plots, below are the details:

Analyze the number of tweets posted over the time

1. The number of tweets has decreased significantly in the last two years
2. The maximum number of tweets, 288 were received in Dec 2015 and the minimum tweets, 34 has been posted in Apr 2017.
3. The line graph shows a constant value of 50 tweets from Apr 2016

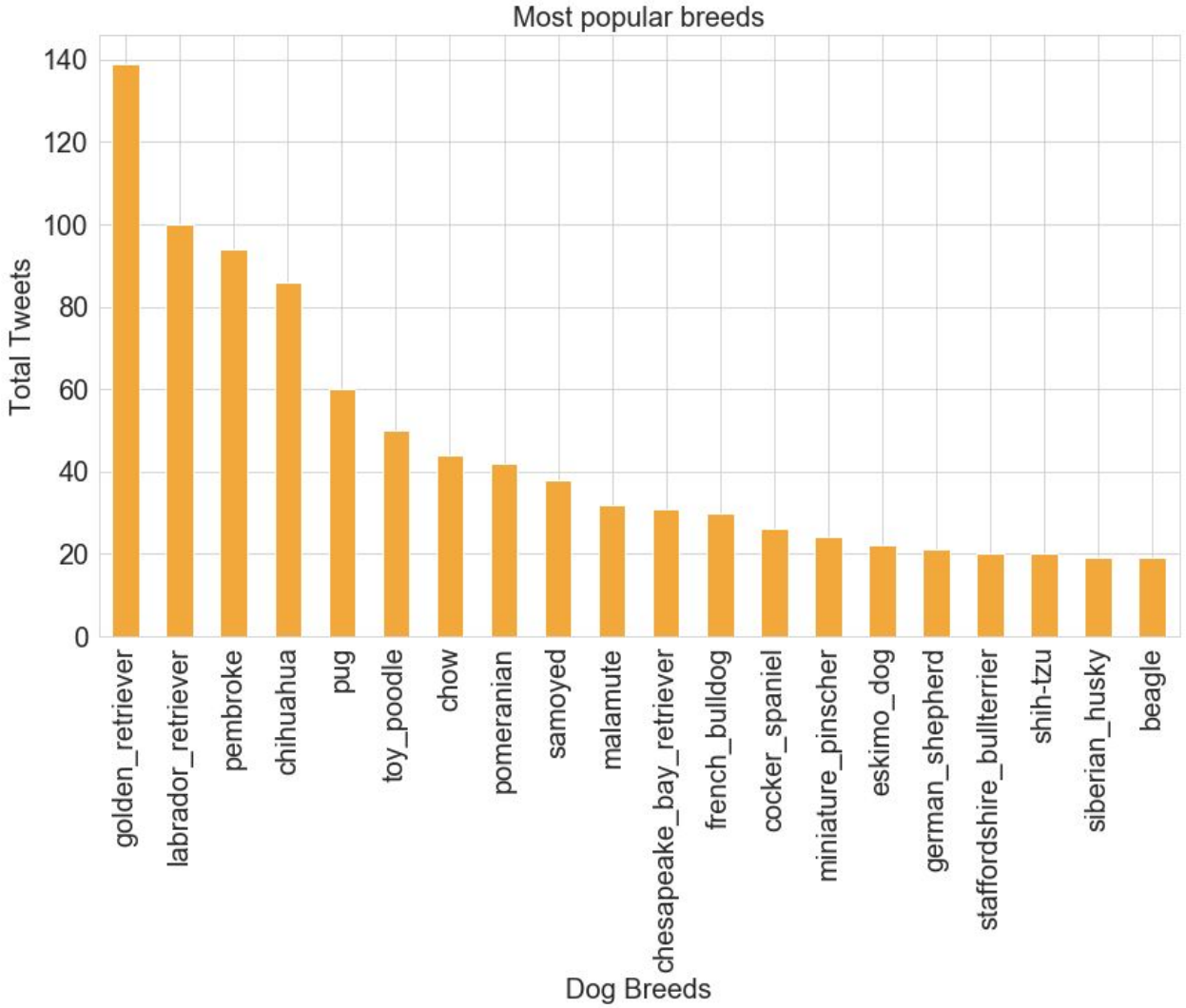


Top 20 popular dog breeds

The top 5 popular dog breeds are:

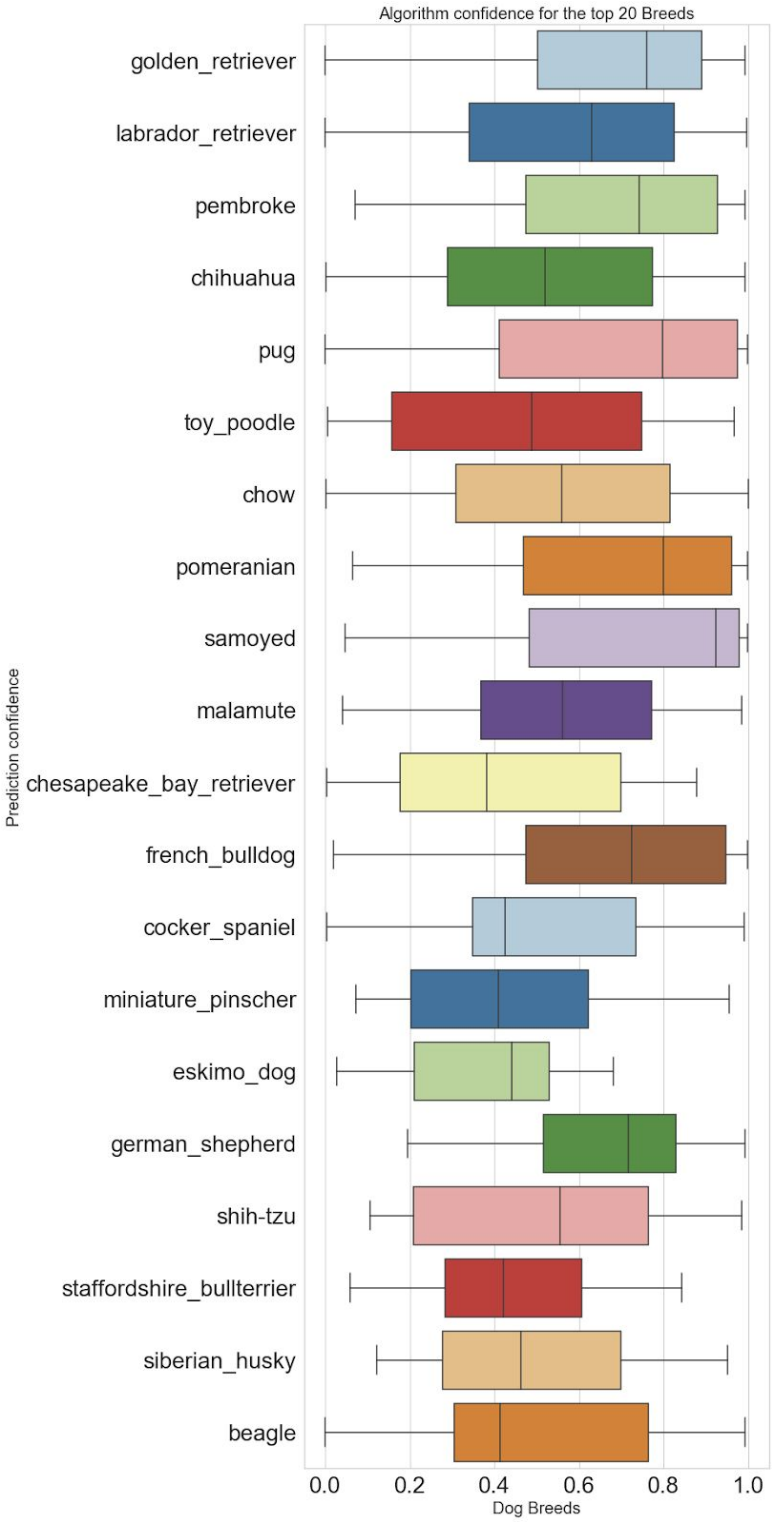
- 'golden_retriever'
- 'labrador_retriever'
- 'pembroke'
- 'chihuahua'
- 'pug'

Bar plot:



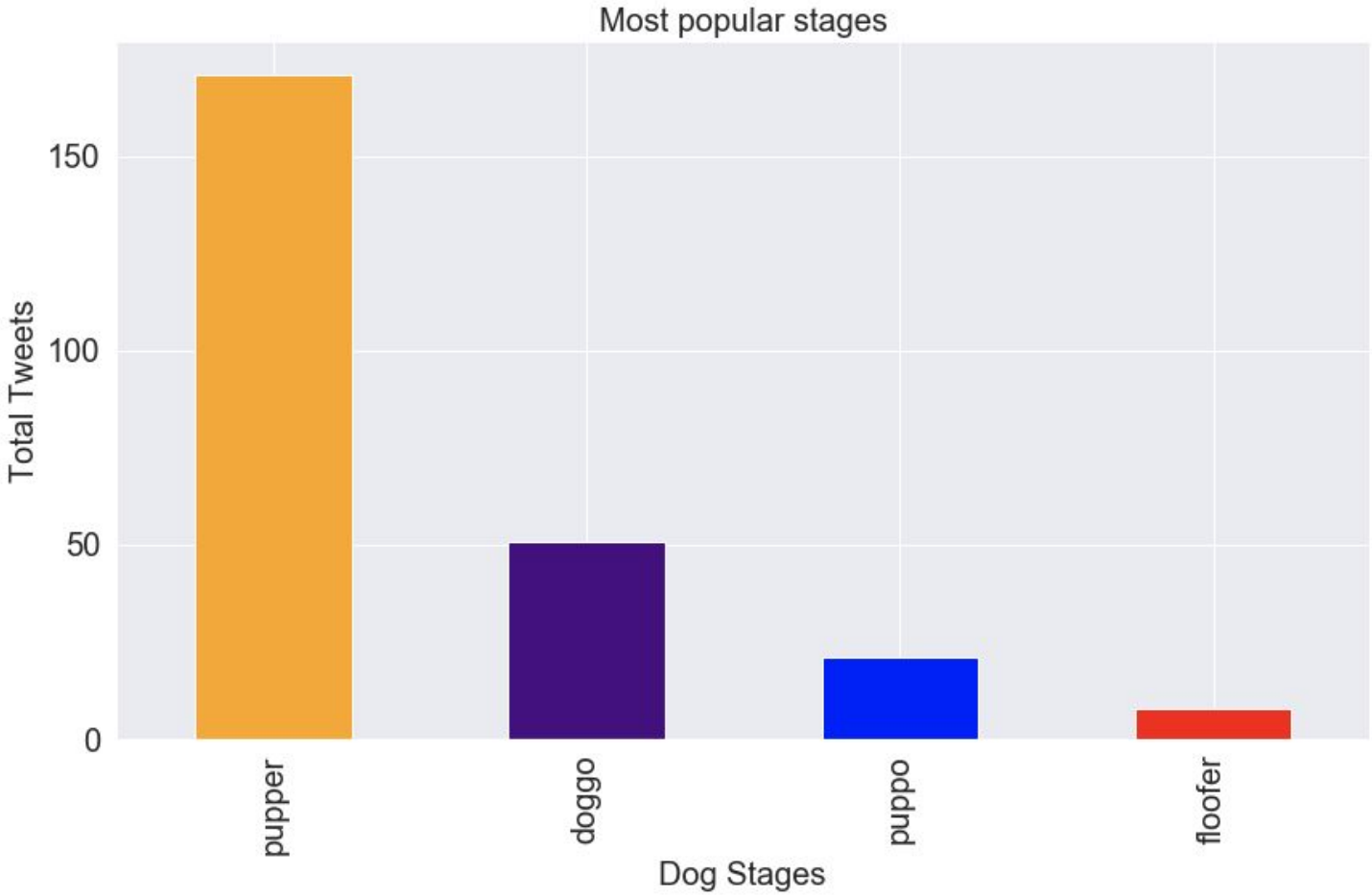
Algorithm confidence for dog breed

- The algorithm prediction works well with the breeds with a higher median. For example - samoyed, pomeranian.
- The algorithm prediction didn't provide a great prediction for breeds like chesapeake_bay_retriever and toy_poodle.



Most popular dog stage

Pupper is the most popular dog stage and floofer is the least.



Analyze the relationship between key variables

- retweet_count and favorite_count has a strong correlation
- no relation between rating denominator and favorite count as the denominator is always 10
- favorite_count has a positive correlation with numerator and dog_confidence but scatter plot will not provide much information