

Project: Wrangle and Analyze WeRateDogs Tweets

Muthukumar Palavesam

Analysis Report

In this document, my insights and visualizations are summarized. It consists of analysis done in python and other libraries.

The motivation for this analysis was to answer the following questions.

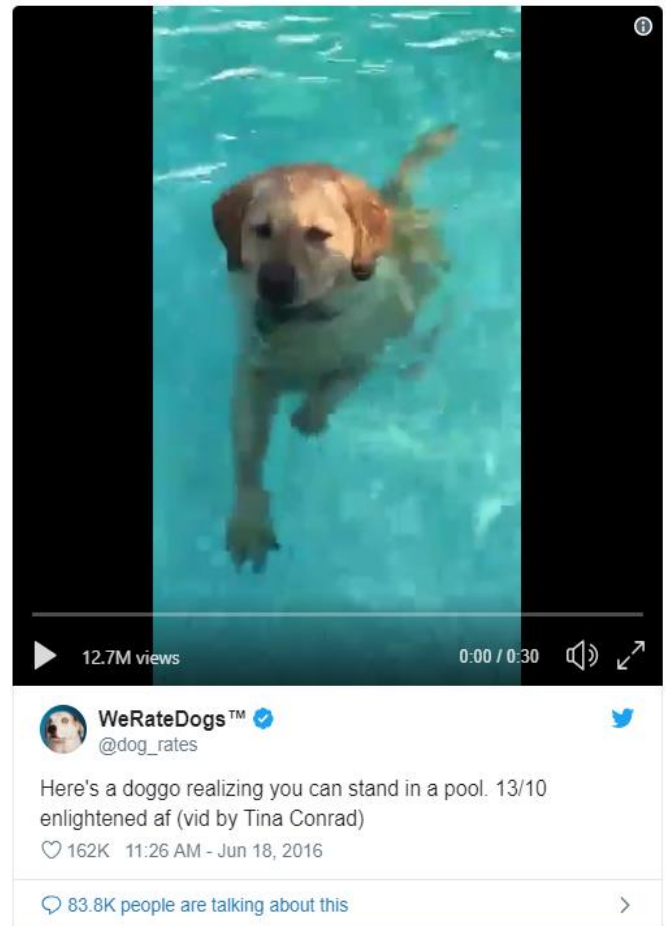
- Which dog breed received the max retweets and likes?
- Which dog breed received the most number of tweets?
- Which Dog breed received more likes in WeRateDogs Twitter feed?
- Top 10 common dog names
- Most used tweet source
- Famous dog Stages
- WeRateDogs tweeting trend over time
- Scatter plot relationship between 'tweets likes count' and 'retweets count'
- Average Retweeting and tweet Likes count trend over time
- Which weekday did people tweet most?
- Which dog breed received the most number of tweets?

Which dog breed received the max retweets and likes?

	tweet_likes_count	retweets_count	rating	breed_confidence
count	1643.000000	1643.000000	1643.000000	1643.000000
mean	8992.383445	2668.727937	10.816799	0.550036
std	13182.668974	4851.563845	1.782965	0.298741
min	77.000000	11.000000	0.000000	0.000010
25%	2032.000000	595.500000	10.000000	0.304538
50%	4212.000000	1329.000000	11.000000	0.550506
75%	11134.000000	3021.500000	12.000000	0.821684
max	162366.000000	82562.000000	14.000000	0.999956

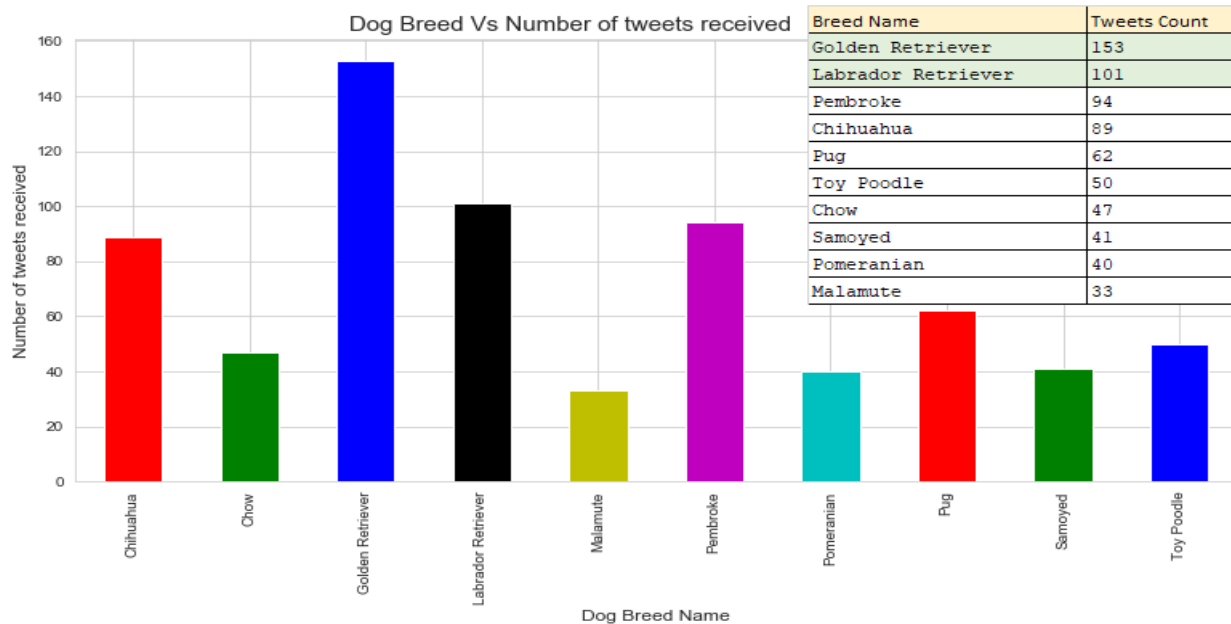
From the above descriptive information results, the max tweet likes count = 162366 and max retweets count = 82562.

tweet_id	744234799360020481
tweet_image_url	https://twitter.com/dog_rates/status/744234799360020481/video/1
tweet_text	Here's a doggo realizing you can stand in a pool. 13/10 enlightened af (vid by Tina Conrad)
tweet_likes_count	162366
retweets_count	82562
rating	13
dog_stage	doggo
dog_breed	Labrador Retriever
breed_confidence	0.825333



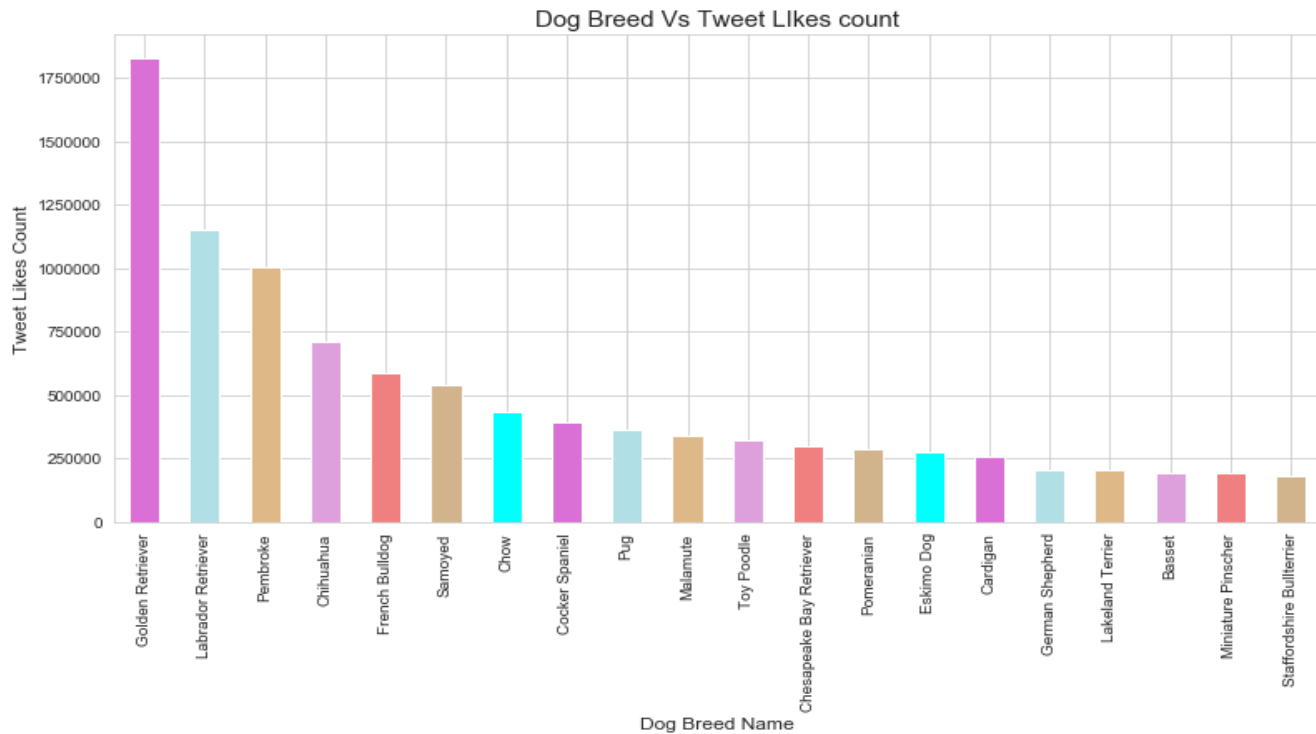
From the descriptive information results, I came to know the max tweet likes count value and max retweets count value. After that, I pulled the exact data and found that the Labrador_retriever doggo swimming and standing in the pool got max retweets and tweet likes

Which dog breed received the most number of tweets?



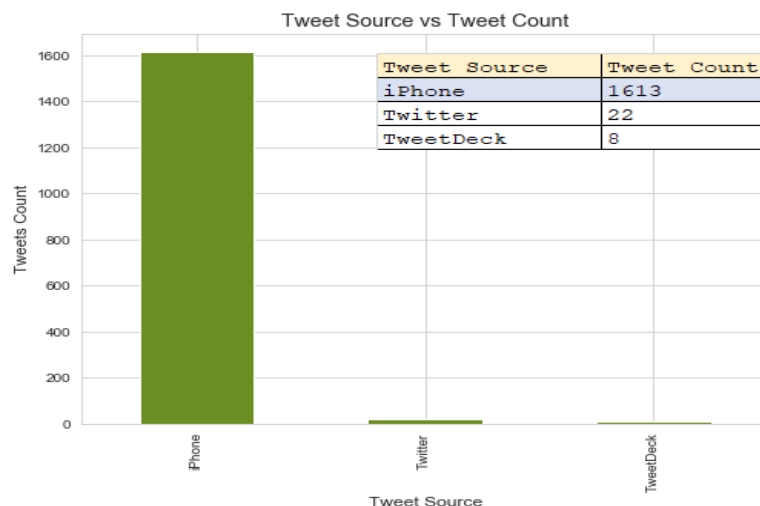
From the above Figure, Golden Retriever and Labrador Retriever had received the most number of tweets with about 150 and 100 tweets, respectively. This suggests that owners of certain dog breeds may use Twitter and send dog pictures to WeRateDogs more often than owners of other dog breeds.

Which Dog breed received the most number of likes in WeRateDogs Twitter feed?



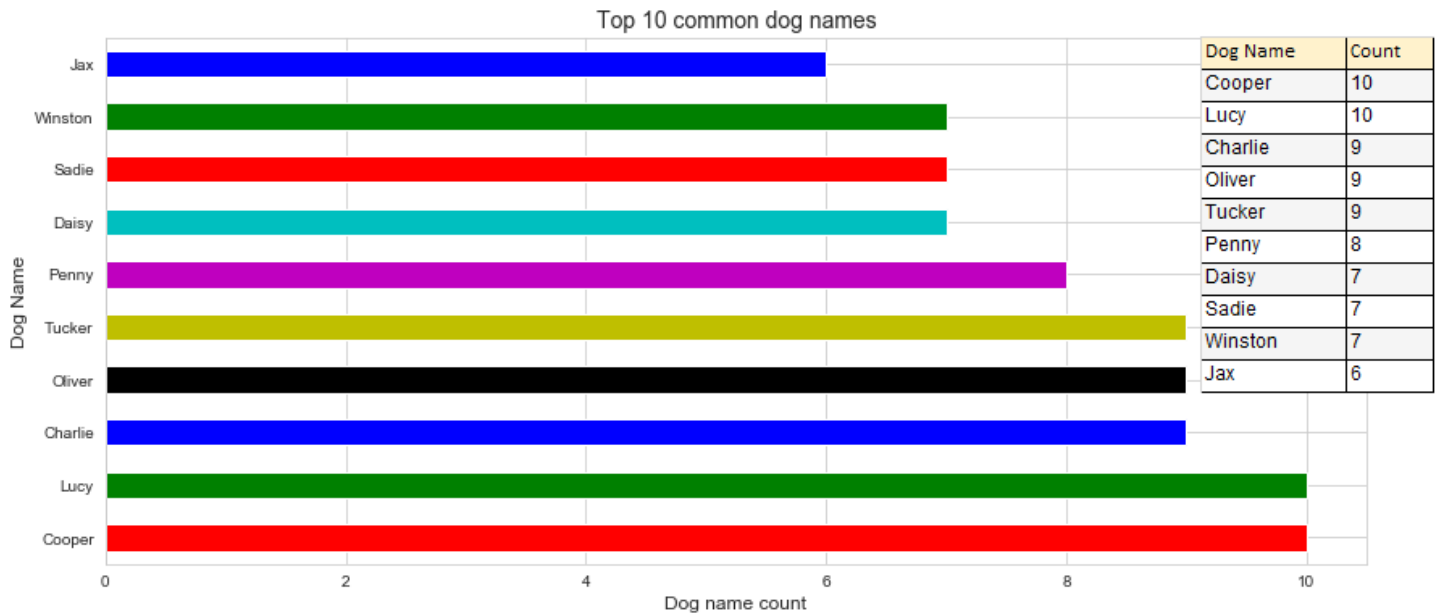
From the above Figure, “Golden Retriever” received the most number of tweet likes and “Labrador Retriever” is the second most number of likes. This suggests that people like the “Golden Retriever” dog breed more compare to other breeds.

Most used twitter sources:



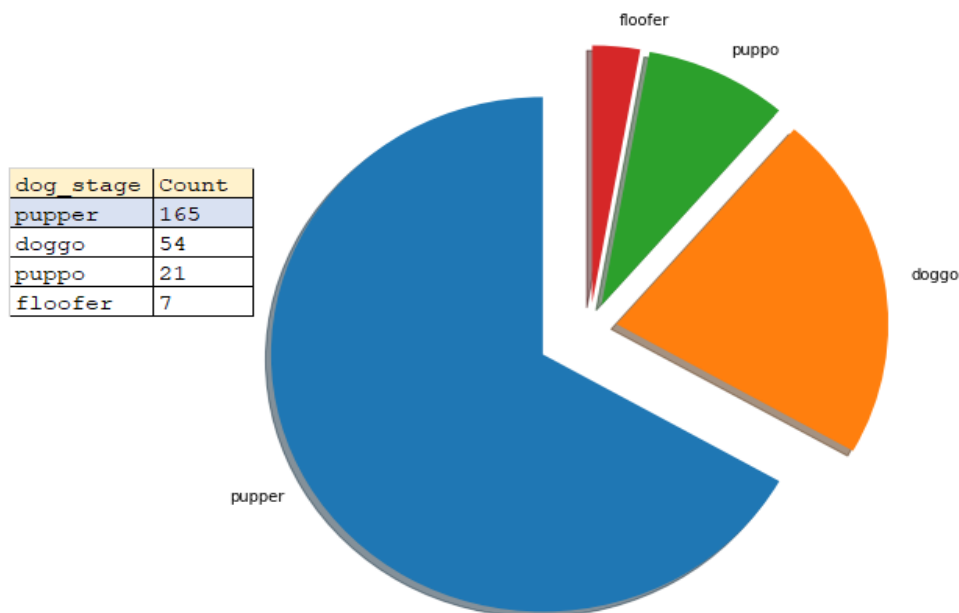
We can see that most of the tweets from iPhone twitter app, which is 98% in the total. Another twitter source is very very less.

Top 10 common dog names:



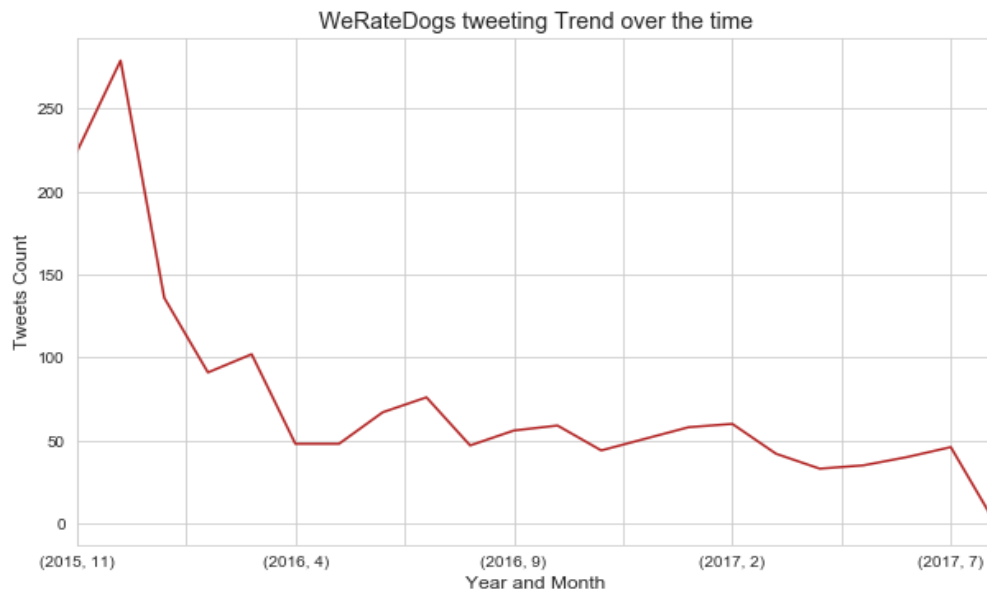
Cooper and Lucy is the most common dog name. 10 of the dogs were named Cooper and Lucy and its the most we have. A close second will be Charlie, Oliver, Tucker Oliver, with 9 dogs being named as these.

Famous dog Stages:



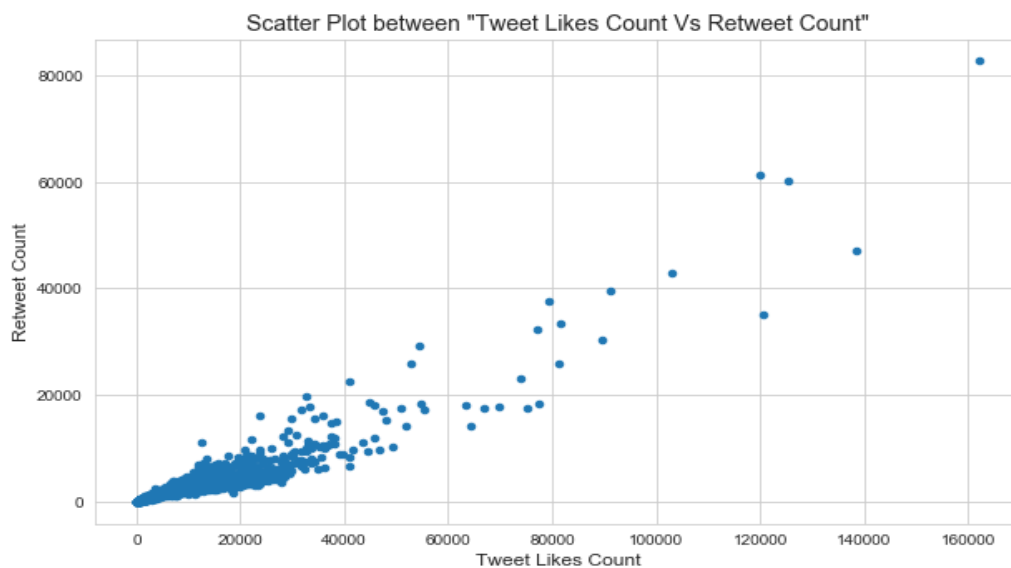
The pie chart is showing as "pupper" is the most popular doggolingo word used in WeRateDogs Tweet page. The number of pupper counts is showing as 165 out of 247

WeRateDogs tweeting trend over time:



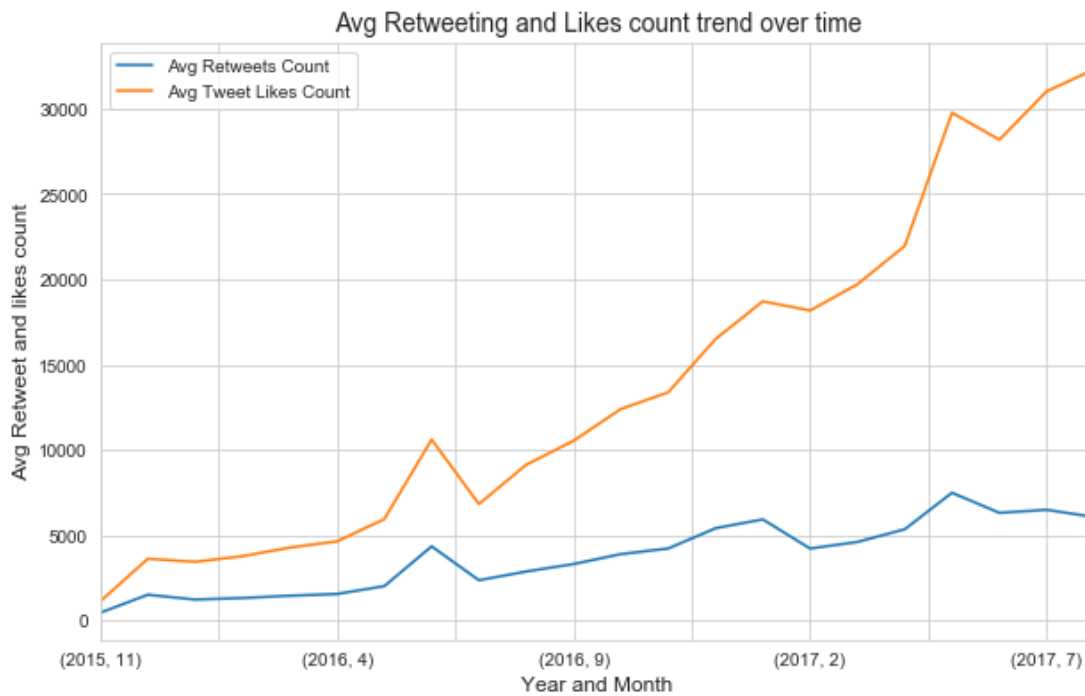
In this analysis, I analyzed WeRateDogs tweets over the time period 11/15/2015 and 8/1/2017. The above line chart clearly showing that the Twitter account started in 2015, tweeted a lot in its starting days, but tweeting activity got reduced drastically.

Scatter plot relationship between 'tweets likes count' and 'retweets count':



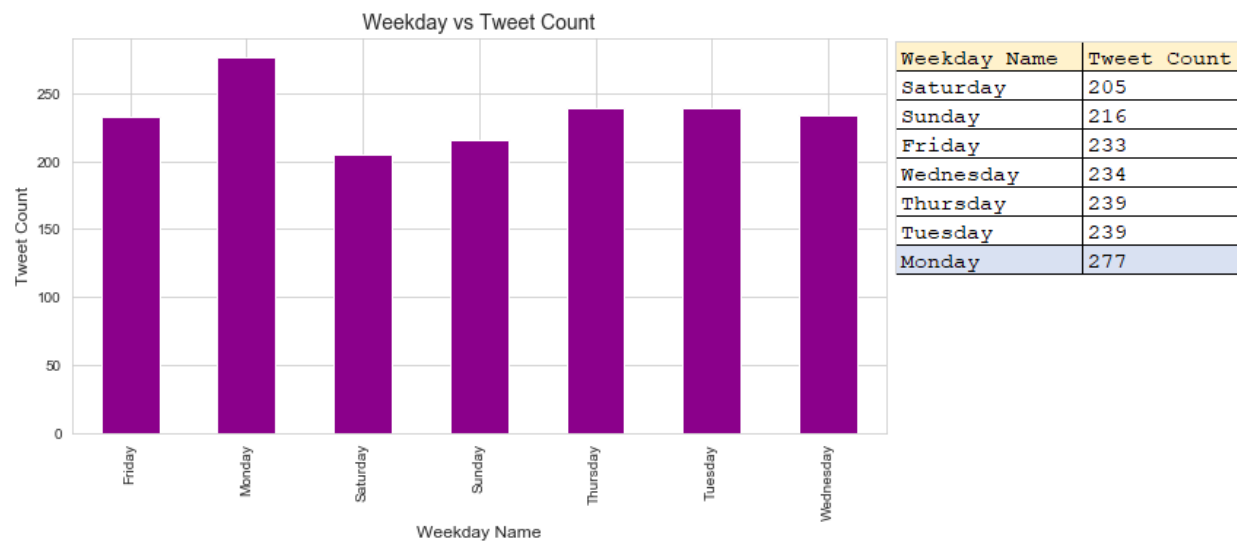
Presented above is a scatter plot of how many tweet likes gets vs retweets. There appears to be a strong positive correlation between "tweet likes count" and "Retweets count". People who favorite seem to almost always retweet.

Average Retweeting and tweet Likes count trend over time:



The above plot shows the "average retweets count" and "Average tweet likes count" increasing since 2015. But if we compare with both, then the "tweet likes count" is higher increase with time.

Which weekday did people tweet most?



The above bar chart clearly showing as that the most number of people tweeted on Monday and Least tweeting on Saturday, Sunday