

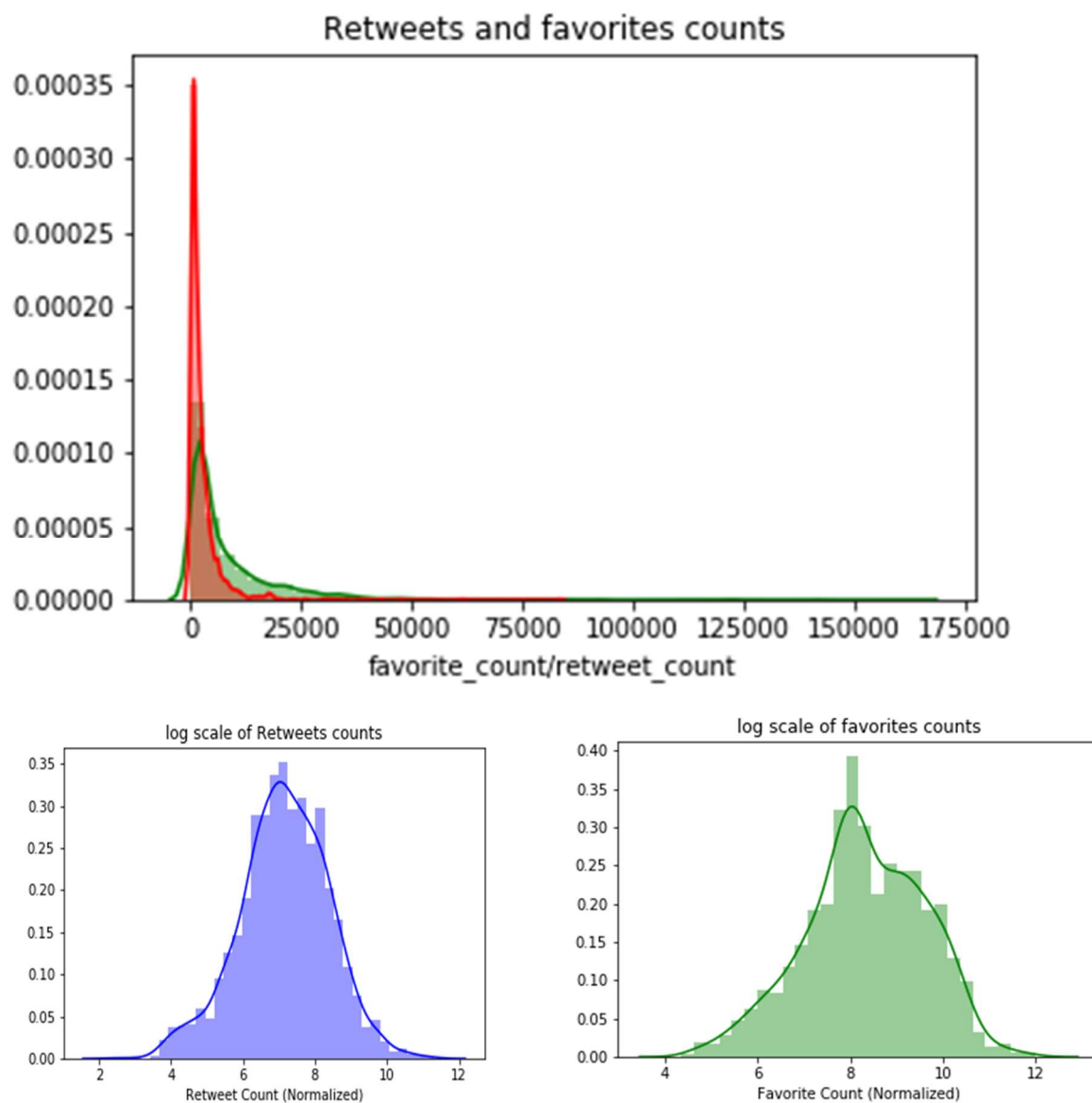
## Introduction:

For this analysis i collected the twitter data (tweet\_count, favourite\_count) using the twitter tweepy api. My analysis revolve around tweet\_count and favourite\_count and their relation over the time period.

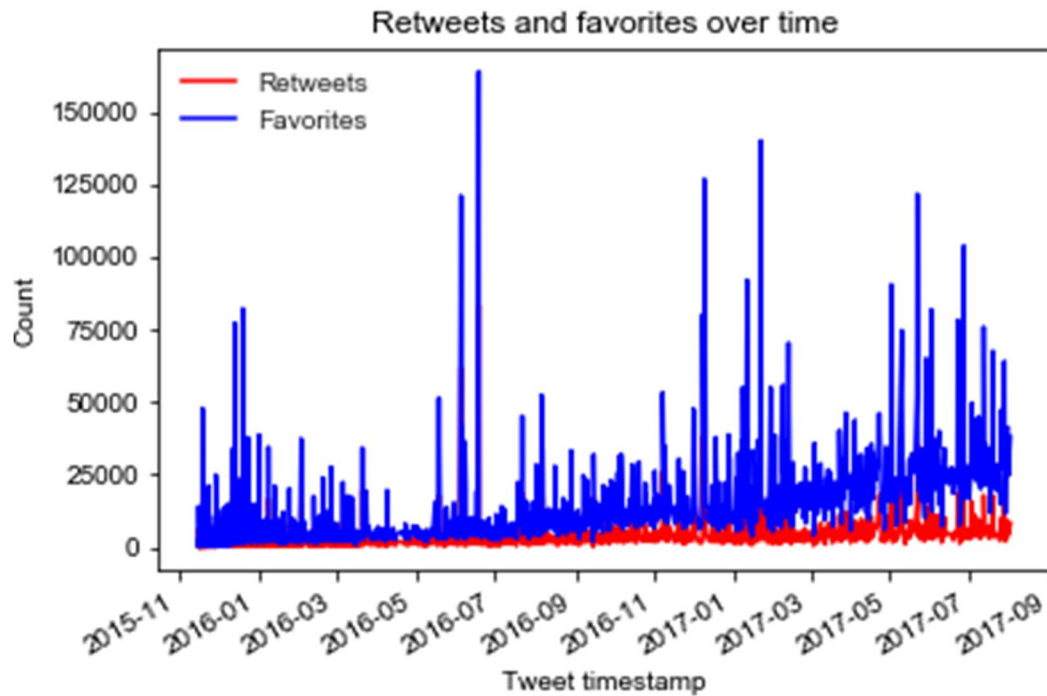
## Analysis:

The original distributions for both favorites and retweets have long positive tails. The normalized graphs again show similar kind of distributions. They are roughly normal except for the spike in values for the normalized favorite count. That may be due to my jitter work and not an actual attribute of the data. Or there are actually a bunch of tweets with one favorite count.

The two variables seem to have similar distributions implying similar phenomena.



In the chart below, we see an upward trend for both retweets and favorites over time. There is a more noticeable increase in the number of favorites when compared to retweets as well as several large outliers in favorites for extremely popular tweets.



There is a strong relationship between retweet and favorite counts (0.92).