ACT REPORT

A Report on the analysis of WeRateDogs Dataset

WeRateDogs is an international, professional twitter account that rates dogs with funny captions or comments. The ratings almost always have a denominator of 10. WeRateDogs account on Twitter has over four million followers. The dataset contains 2356 tweets of the 5000 tweets between 2015 and 2017.
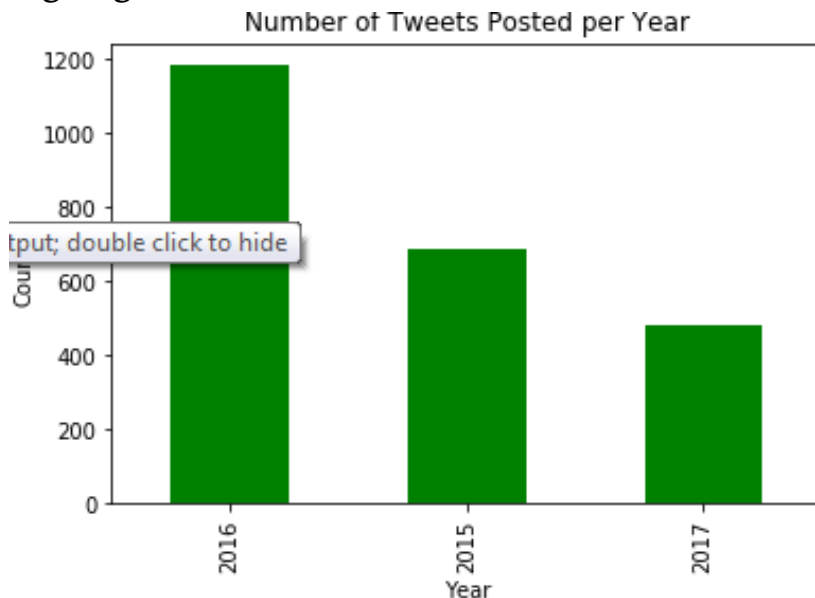
The data set was gathered from three sources:

1. The Enhanced Twitter Archive, downloaded manually, as a CSV (Comma Separated Values) file from the Udacity portal.
2. The Tweet Image Prediction, a TSV (Tab Separated Values) file, downloaded programmatically via a link provided by Udacity
3. The Tweet Json file, gotten from the Udacity website. I could not get a Twitter developer account, so I used the alternative provided by Udacity

Data Wrangling: The data was gathered, accessed, cleaned, and tested. These steps were explained in the data wrangling report
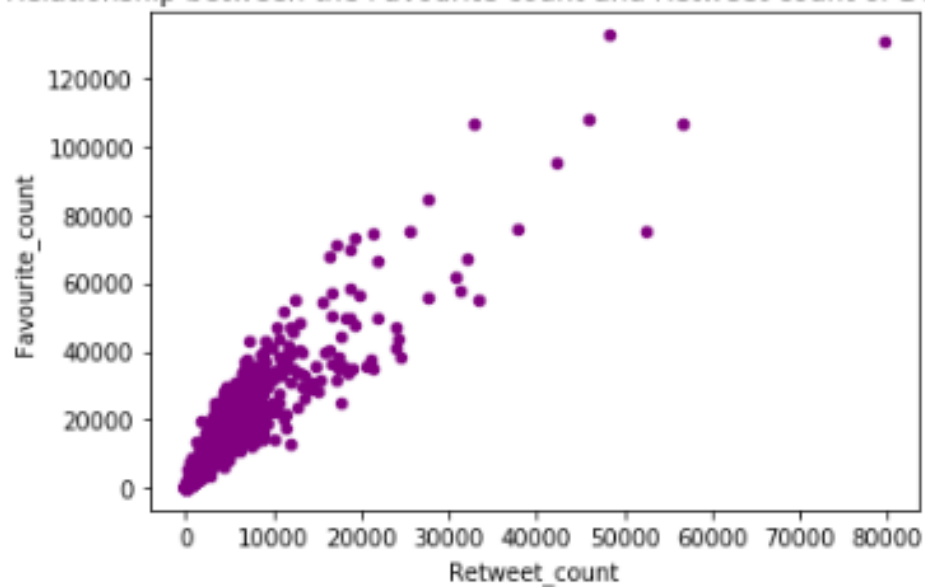
The Insights gotten form the data are:

1. Charlie and Lucy, are the most common dog names with 11 appearances each.
2. The dog with the highest rating is Pumpkin. It is a Greater Swiss Mountain Dog. It's dog stage is not known.

Number of Tweets Posted per Year

Relationship between the Favourite count and Retweet count of Dog Tweets

4.   The number of retweets have a positive correlation with the favourite count of each tweet.