We Rate Dogs Project

Act report

Mohamed Makki

January 10th, 2020

Today, I would like to share with you some interesting findings extracted from data collected by Twitter; the resulting information and all insights and visualizations obtained after our data has carefully wrangled and cleaned. This topic is pertaining to over 5000 of tweets and data gathered by Twitter's from followers from all over the world under attractive title "WeRateDogs".

It is obvious from the titles that twitter solicited the users/followers to vote (i.e like) and pose opinions for rating various breeds of dogs, the event have taken place during the period from Nov. 15th, 2015 till Aug 1st , 2017.


After duly wrangling and cleaning the data, the conclusion was interesting about what the collected data could tell us about these contributions.

The visualizations show that most of the tweets counts lies within rating values 10 to 12 (or as per twitter scale of rating = 10/10 to 12/10).

The total number of dog breed were 113 breed of dog in the prediction column

Distribution of Tweets on Dogs Rating
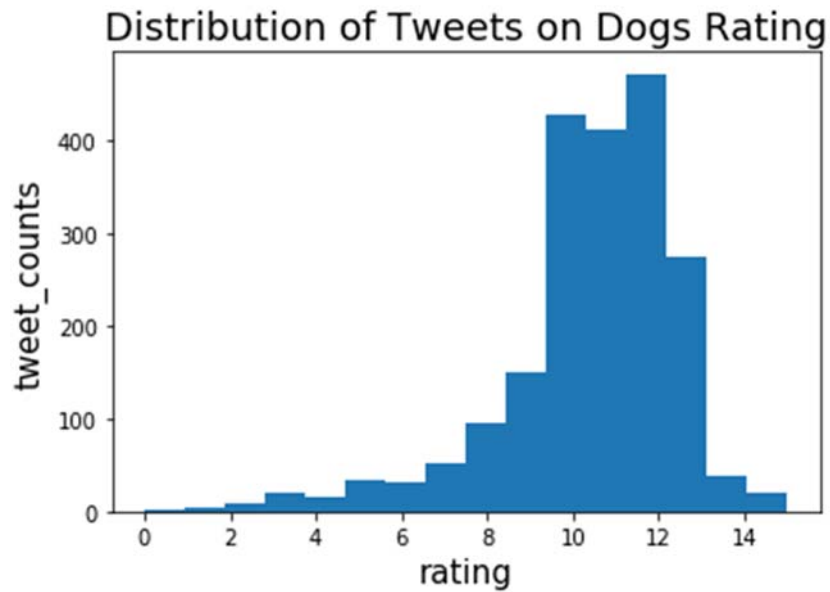
In the other hand, we find that there is a strong and positive relationship (correlation coefficient) between all retweet counts and all likes counts. In same direction, there is the same high connection (correlation) between each dog_stage retweet counts and their likes (each of dog_stage separately).

Likes Distribution Over The Retweet count
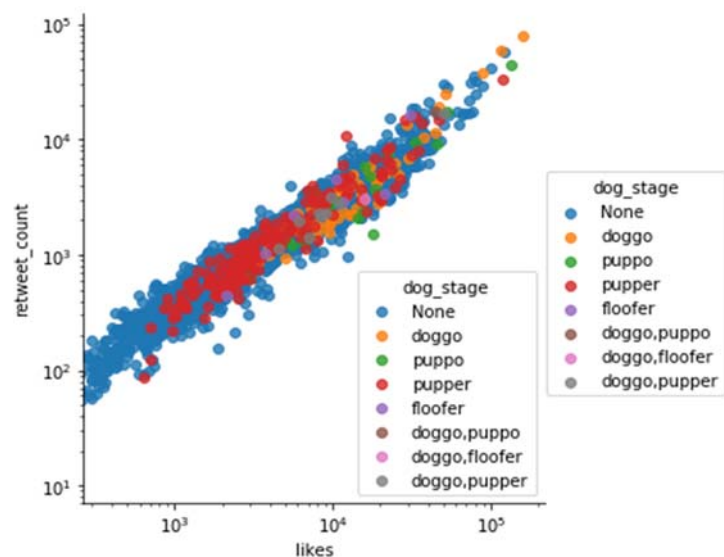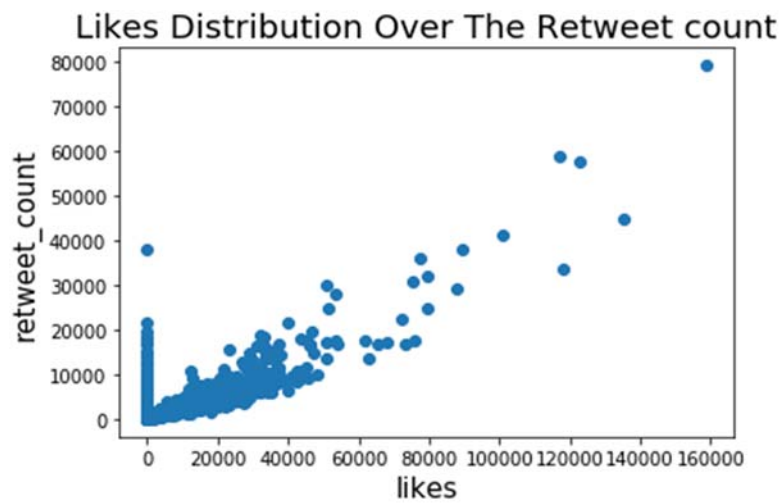
Regarding the distribution of dog stage in this rating, (most available stage of dogs) will discover that most stages were unknown (about 1742 of 2061 = 85%) but for the known stages the following table will show us the distribution of dog_stage in Twitter dog rating event.

No.	Dog Stage	Counts	Percentage
1	Pupper	211	66%
2	Doggo	65	20%
3	Puppo	23	7%
4	Doggo,Pupper	11	4%
5	Floofer	7	2%
6	Doggo,Puppo	1	0.5%
7	Doggo,Floofer	1	0.5%