Report: act_report

Insights from the cleaned data

Having cleaned the various datasets and merging them into a singly master dataset, *twitter_archive_master.csv*, we did some simple analysis on the master dataset. Below are some insights that were gathered from the analysis done.

From the image below, the least favorite dog got a rating of 2. About 75% of dogs in the data had a favorable rating of 12. About 50% of the dogs had a rating of 11

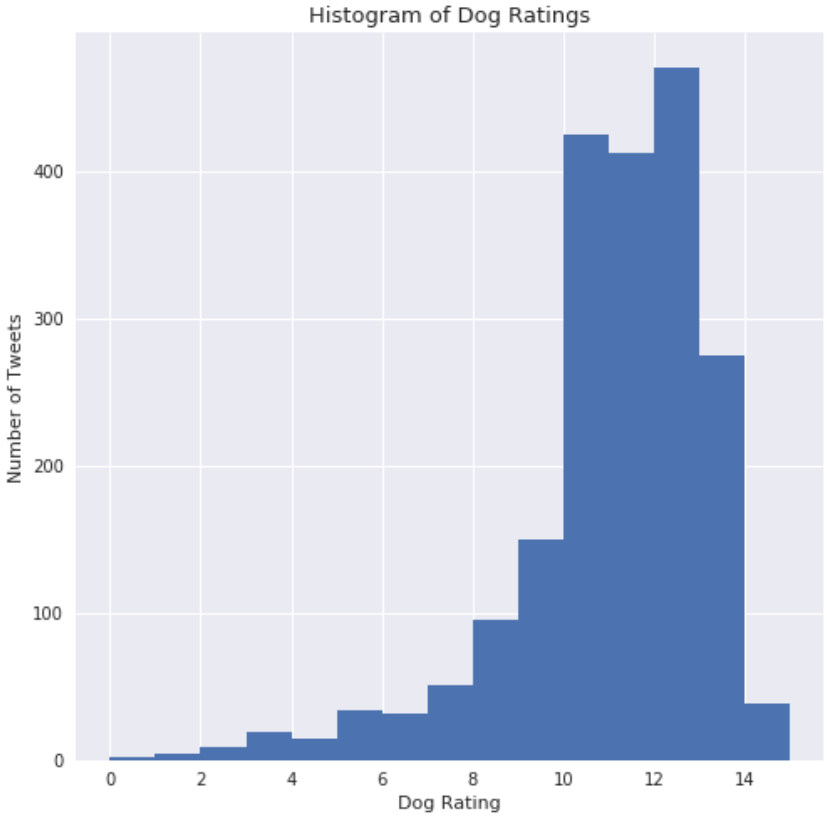
	rating_numerator	retweet_count	favorite_count
count	1172.000000	1172.000000	1172.000000
mean	10.925768	2216.180034	8009.719283
std	1.684952	3867.743845	11538.098747
min	2.000000	11.000000	66.000000
25%	10.000000	515.250000	1859.500000
50%	11.000000	1123.500000	3800.500000
75%	12.000000	2569.000000	9938.750000
max	14.000000	52505.000000	123439.000000

It was observed that on average, dogs got a rating of approximately 11. The most favorite dogs got ratings of 14. Below is a screenshot of sample tweets of the most favorite dogs with ratings of 14:

tweet_id	text	rating_numerator	name	dog_stage	retweet_count
890240255349198849	This is Cassie. She is a college pup. Studying international doggo communication and stick theory. 14/10 so elegant much sophisticate https://t.co/t1bfwz5S2A	14	Cassie	doggo	6053
884441805382717440	I present to you, Pup in Hat. Pup in Hat is great for all occasions. Extremely versatile. Compact as h*ck. 14/10 (IG: itselizabethgales) https://t.co/vvBOcC2VdC	14	None	None	4769

Cassie seems to be a sweet little pup with some very fancy course of study for doggos. It isn't surprising that she got a rating of 14 and a retweet count of 6,053. See how cute she looks in in her glasses in the picture below:

Overall, there is a generally positive rating for dogs during the period under review. This is shown in the graph below:



Observations from the master dataset also shows that the highest retweet of an original tweet about a dog during the period under consideration was 70,334 retweets, and the highest favorite count about a tweet was 144,245.

Again, it was observed from the dataset that Twitter for iPhone was the most predominant display type/source for tweets. This accounted for about 1,154 of the tweets. This is shown in the screenshot and pie chart below:

```
#the most common display source
df.source.value_counts()
```

```
Twitter for iPhone     1154
Twitter Web Client       14
TweetDeck                 4
```