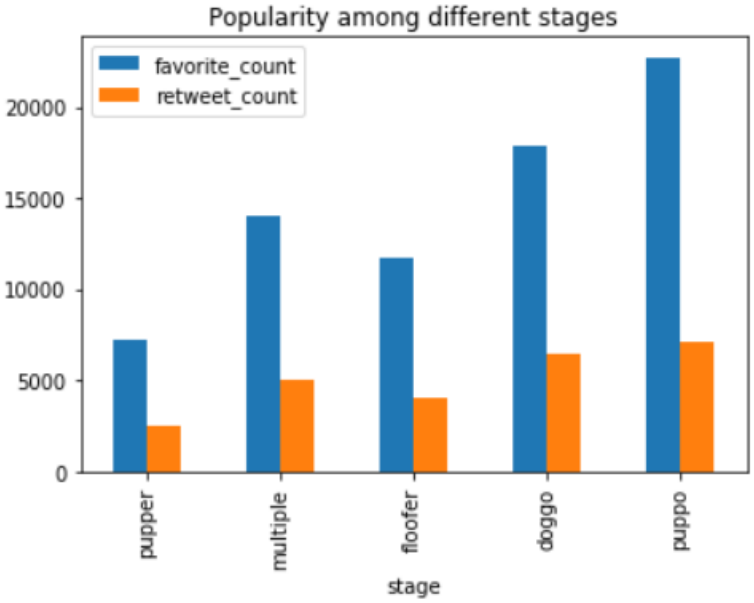
Using the master data that is assessed and cleaned and below insights were found

1. **Popularity among different stages**
   Goal: To find out which stage is most popular.
   Data: stage, rating, favorite_count and retweet_count columns
   Assess: For each stage, the mean of rating, favorite_count and retweet_count are calculated
   Result:

stage	rating	favorite_count	retweet_count	count
pupper	10.739955	7257.527027	2485.837838	222
multiple	10.833333	14017.333333	5010.833333	6
none	11.769336	8605.789324	2659.643384	1761
floofer	11.800000	11674.900000	4083.600000	10
doggo	11.837838	17850.945946	6507.986486	74
puppo	12.041667	22715.125000	7124.875000	24

   Plot:



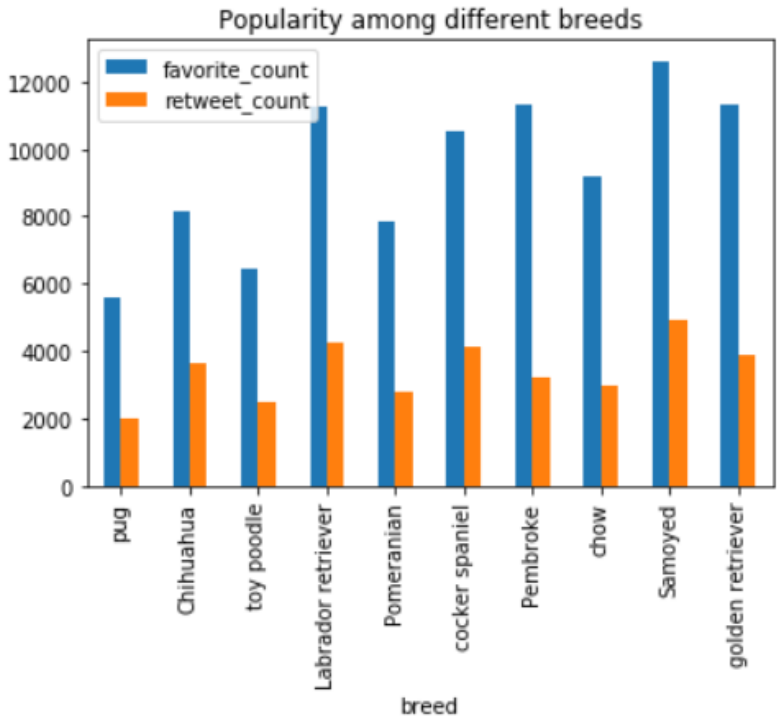
   Analysis:
   From the results, we can see that puppo has the highest favorite_count and highest rating, however, we have only 10 rows of data for floofer. Also, for most of the rows we do not have the stage data.

2. **Popularity among different breeds**
   Goal: To find out which breed is most popular.

Data: breed, rating, favorite_count and retweet_count
Assess: For top 10 breeds, the mean of rating, favorite_count and retweet_count are calculated
Result:

breed	rating	favorite_count	retweet_count	count
pug	10.240741	5617.947368	1979.473684	57
Chihuahua	10.692308	8149.891566	3665.277108	83
toy poodle	11.054054	6472.974359	2472.256410	39
Labrador retriever	11.127660	11238.500000	4273.430000	100
Pomeranian	11.151316	7871.421053	2803.105263	38
cocker spaniel	11.296296	10527.500000	4097.966667	30
Pembroke	11.443182	11304.034091	3185.988636	88
chow	11.609756	9187.295455	2956.363636	44
Samoyed	11.700000	12627.697674	4940.418605	43
golden retriever	11.808686	11310.846667	3874.466667	150
none	13.347003	6444.122722	3481.766707	823

Plot:



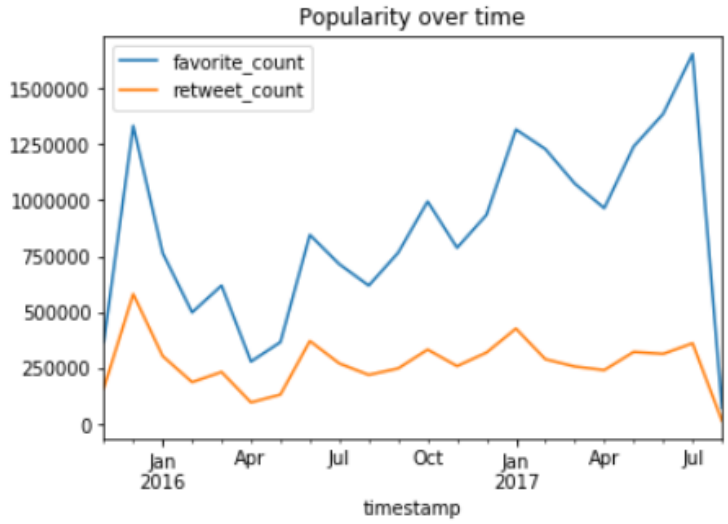
Analysis:

From the results we can see that Samoyed has the highest favorite_count as well as retweet_count even though it is in second place in terms of rating. And Golden retriever is second in place for favourite_count and retweet_count, but first in rating.

3. **Popularity over time**

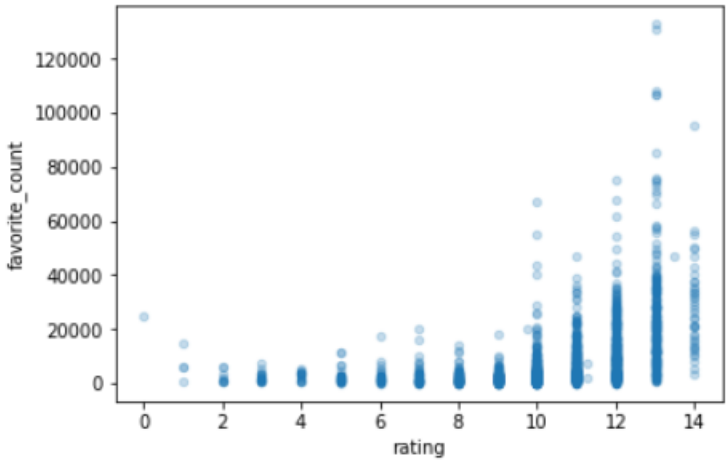
Goal: To find the trend of favorite_count and retweet_count over time
Data: timestamp, favorite_count and retweet_count columns
Assess: favorite_count and retweet_count columns are aggregated for every month
Result:

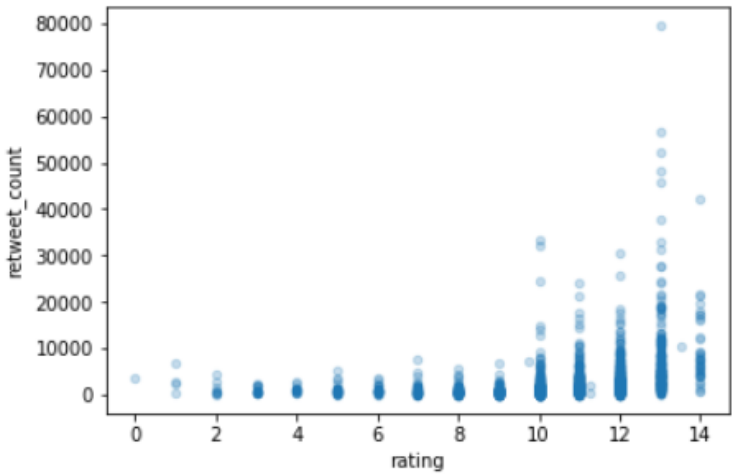
timestamp	favorite_count	retweet_count
2015-11-30	367571.0	160952.0
2015-12-31	1332903.0	581632.0
2016-01-31	764848.0	304553.0
2016-02-29	499740.0	188185.0
2016-03-31	619372.0	233415.0
2016-04-30	280103.0	97674.0
2016-05-31	366741.0	132613.0
2016-06-30	845725.0	371879.0
2016-07-31	714892.0	271731.0
2016-08-31	619401.0	220630.0
2016-09-30	766159.0	249528.0
2016-10-31	994975.0	333664.0
2016-11-30	788055.0	259787.0
2016-12-31	935081.0	320273.0
2017-01-31	1316207.0	427323.0
2017-02-28	1228776.0	290618.0
2017-03-31	1074460.0	257986.0
2017-04-30	964846.0	242027.0
2017-05-31	1240533.0	323109.0
2017-06-30	1385809.0	314276.0
2017-07-31	1653469.0	361755.0
2017-08-31	73286.0	15367.0

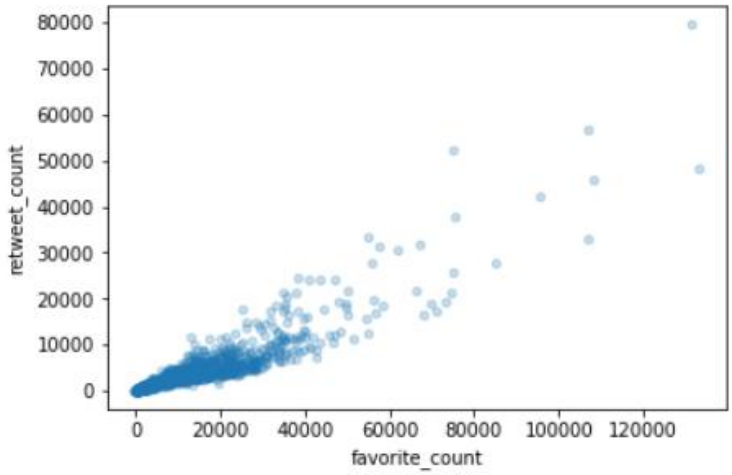


Analysis:
From the plot we could see that it was popular around December 2015 and there was a drop in the popularity and then it gradually increased over time.

4. **Correlation between rating and favorite_count and retweet_count**
   Goal: to find the correlation between rating and favorite_count and retweet_count
   Data: rating, favorite_count and retweet_count columns
   Plot:







Analysis:
We could see that there is a strong positive correlation between favorite_count and retweet_count from the plot. And rating and favorite_count/retweet_count are correlated as well.