The data frame contains 1991 observations. No retweets data.

All observations have image url in the prediction data frame.

Introduction to the data set:

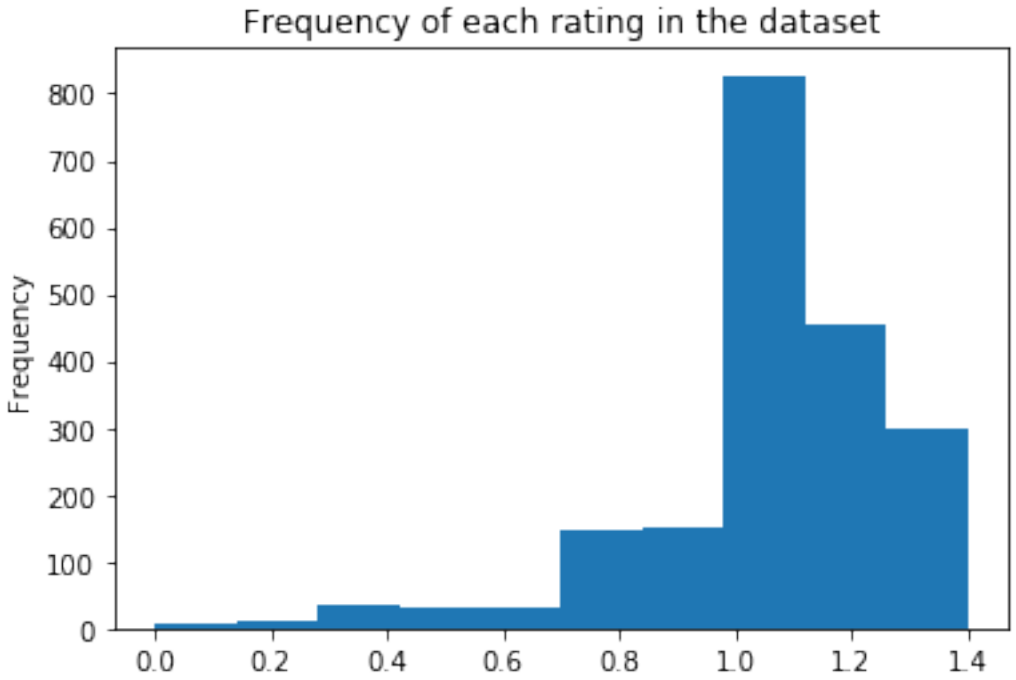
```
Data columns (total 13 columns):
tweet_id            1991 non-null object
source              1991 non-null object
text                1991 non-null object
expanded_urls       1991 non-null object
name                1376 non-null object
gender              860 non-null object
rating              1991 non-null float64
favorite_count      1991 non-null int64
retweet_count       1991 non-null int64
tweet_year          1991 non-null object
tweet_month         1991 non-null object
tweet_day           1991 non-null object
dog_stage           306 non-null category
```

Some Basic Statistics about numeric values:

	rating	favorite_count	retweet_count
count	1991.000000	1991.000000	1991.000000
mean	1.054495	8815.177800	2697.127574
std	0.219063	12655.468769	4718.815468
min	0.000000	80.000000	12.000000
25%	1.000000	1902.000000	601.000000
50%	1.100000	4001.000000	1296.000000
75%	1.200000	11063.500000	3101.000000
max	1.400000	142312.000000	76674.000000

We find from the above statistical data that most of the observation (75%) have rating 1.2 or more.

Let's have a lock to the frequency of ratings:

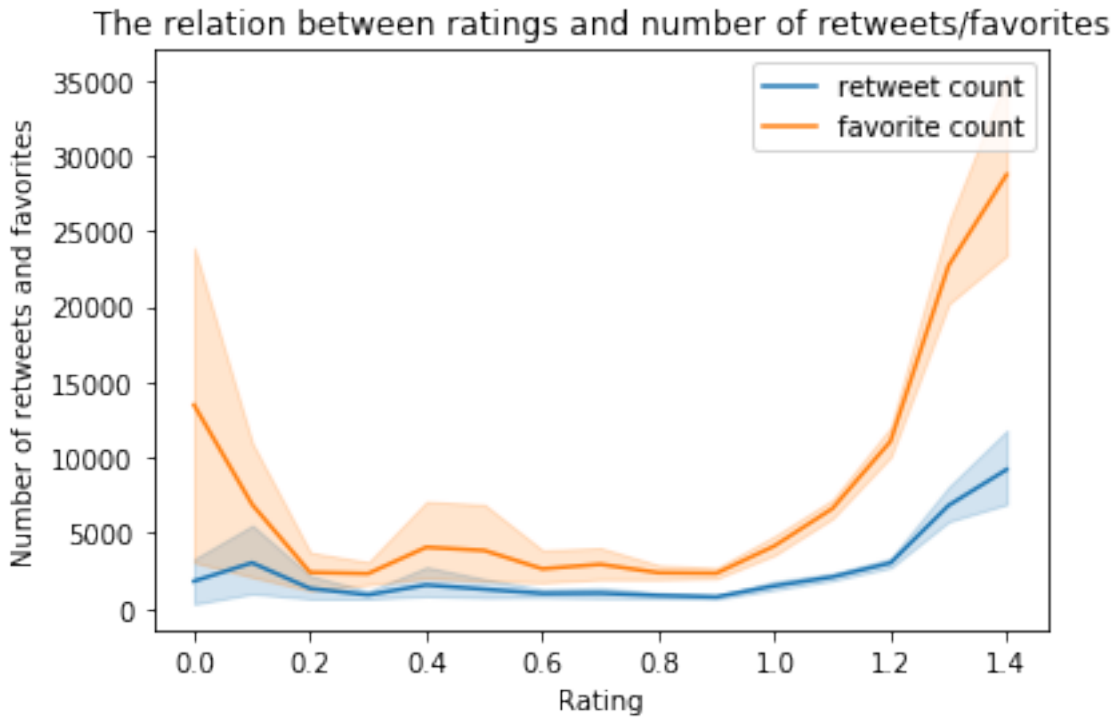


Most of the ratings are between 1.0 and 1.4. Count of ratings below 1.0 is less than ratings more than 1.0

What about twitter data source?

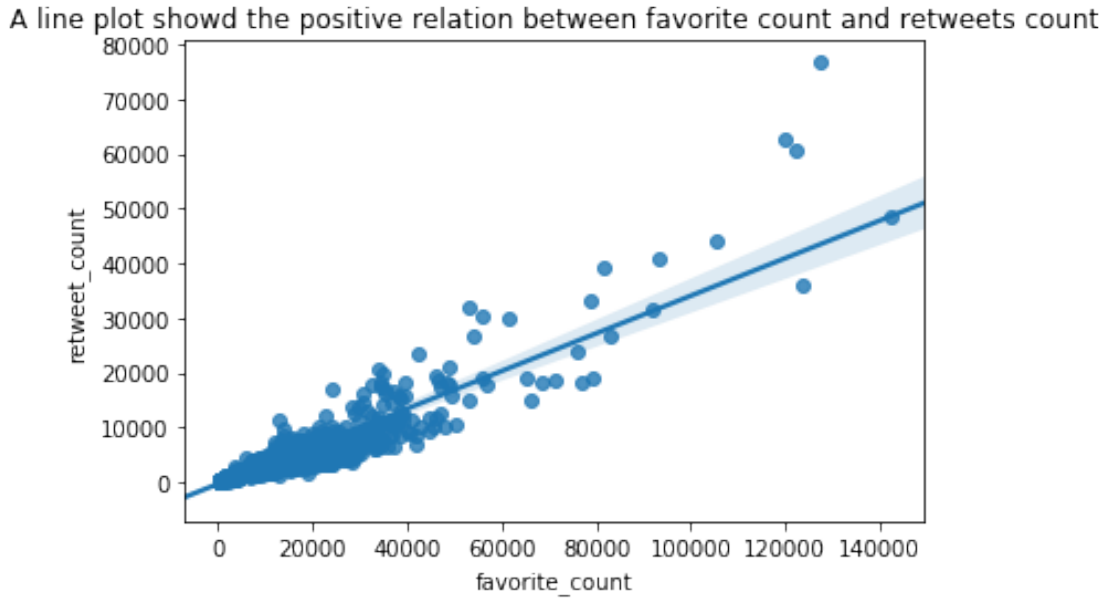
```
Twitter for iPhone    1953
Twitter Web Client      28
TweetDeck               10
```

Relation between retweets and favorites with the rating.

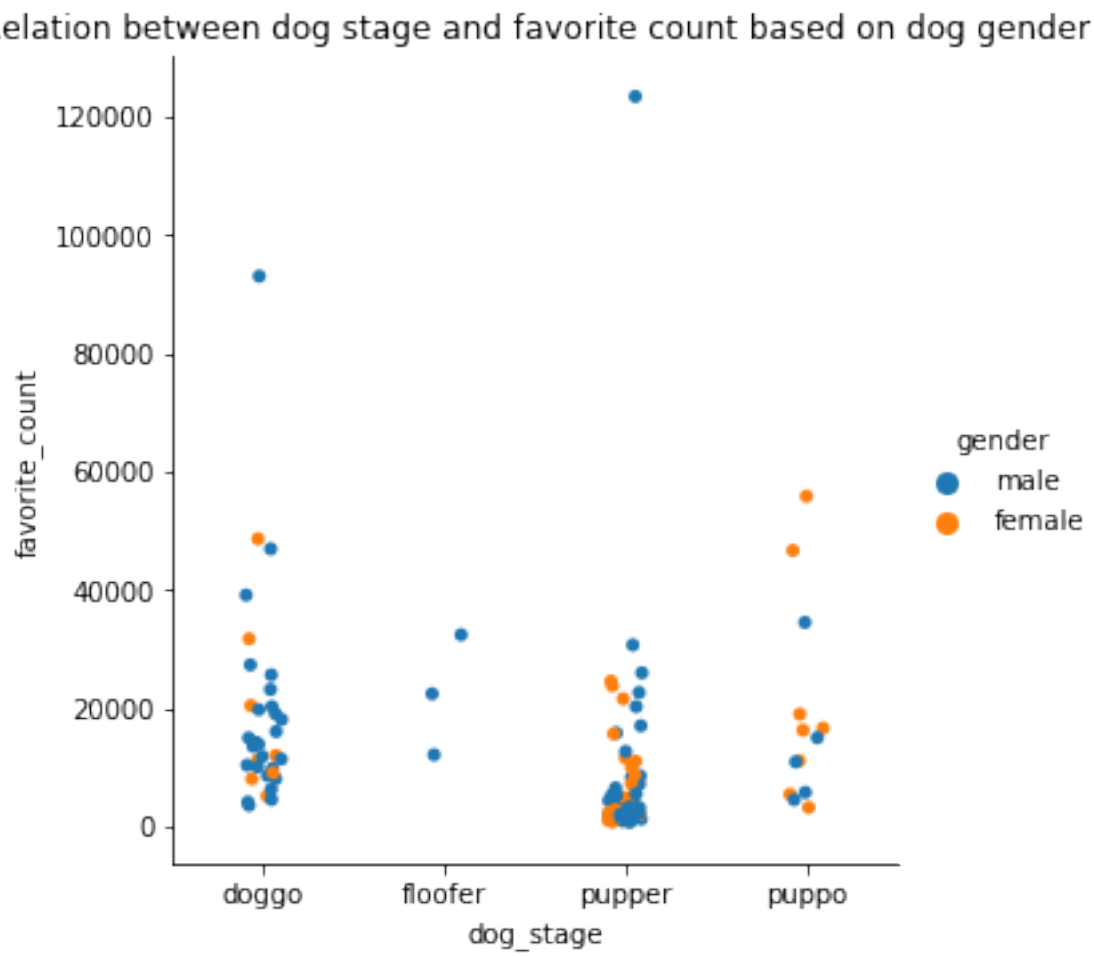


Retweet count are always more than favorite count. Specially ratings over 1.0.

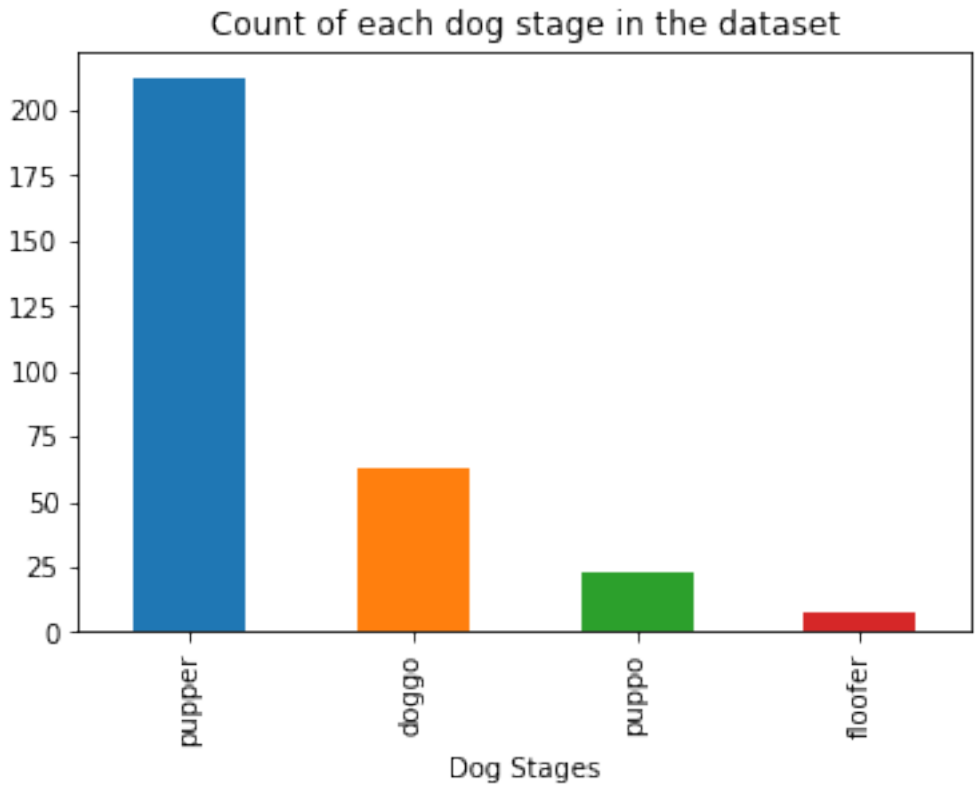
Sounds interesting, let's have a closer look to the relation between retweets and favourits.



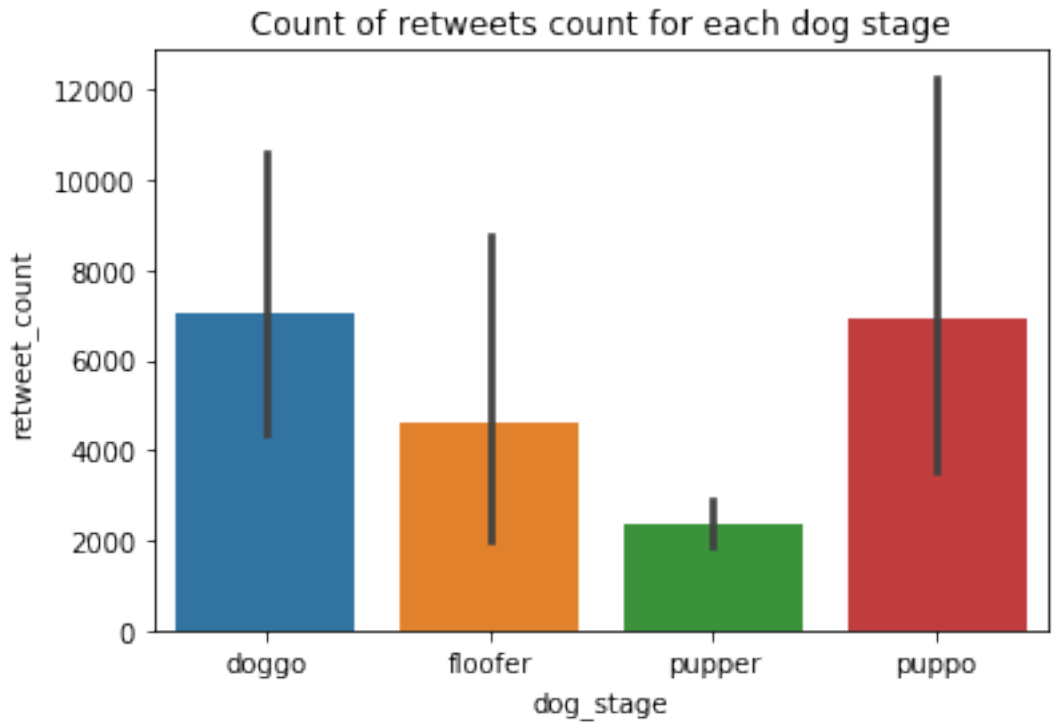

```
Correlation is 0.9220463274686362.
There is a strong relation between retweets and likes.
```



Relation between dog stage and favorite count based on gender.



According to the above plot, we see that 'pupper' have the top values in our data set. **Let's have a closer look to the popularity between dog stages. Which stage has the more retweets?**



'puppo' wins in this comparison. However 'pupper' has the most top frequency.

Let's check the mean rating for the different dog stages:

```
doggo      1.188889
floofer    1.187500
pupper     1.065094
puppo      1.204348
```

Conclusion:

Based on analytics made on We Rate dogs twitter archived data, it's better to buy puppo.

People love puppo and doggo stage more than floofer and pupper stages. We rate dogs also suggest to buy puppo more than other three stages.

Puppers represent the big number of our data, but it has the lowest mean rating.