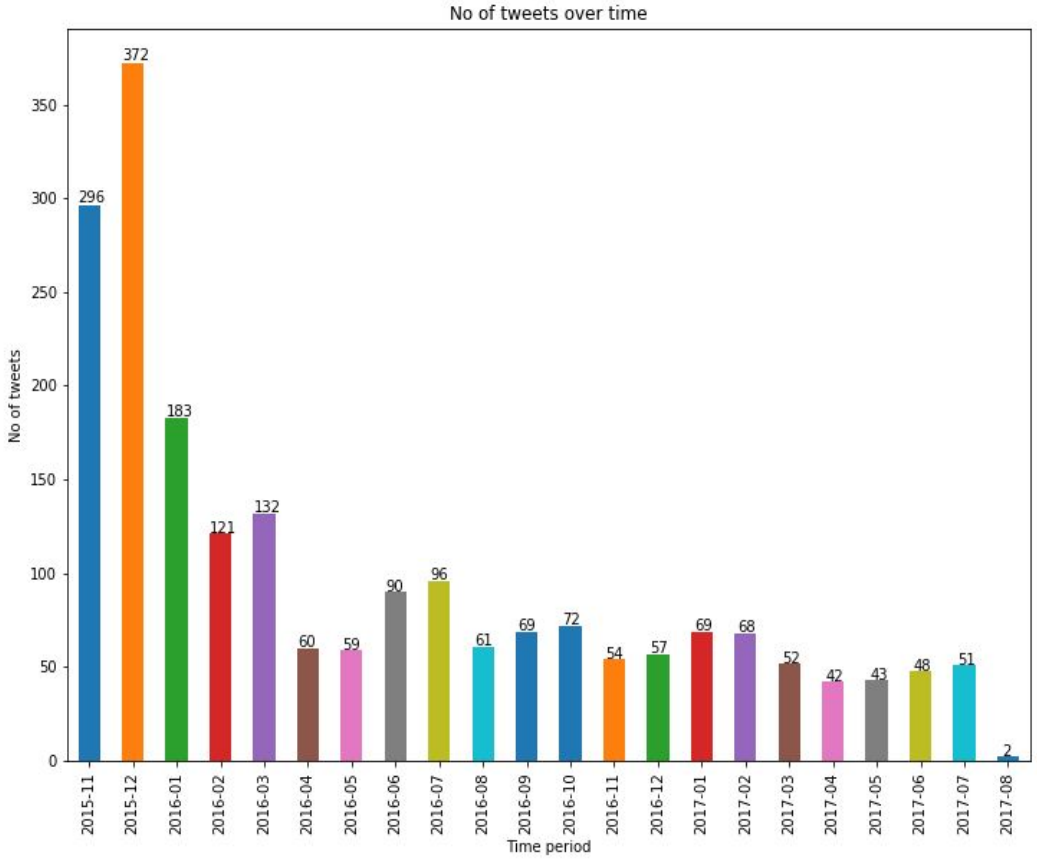
Data Analysis and visualisation

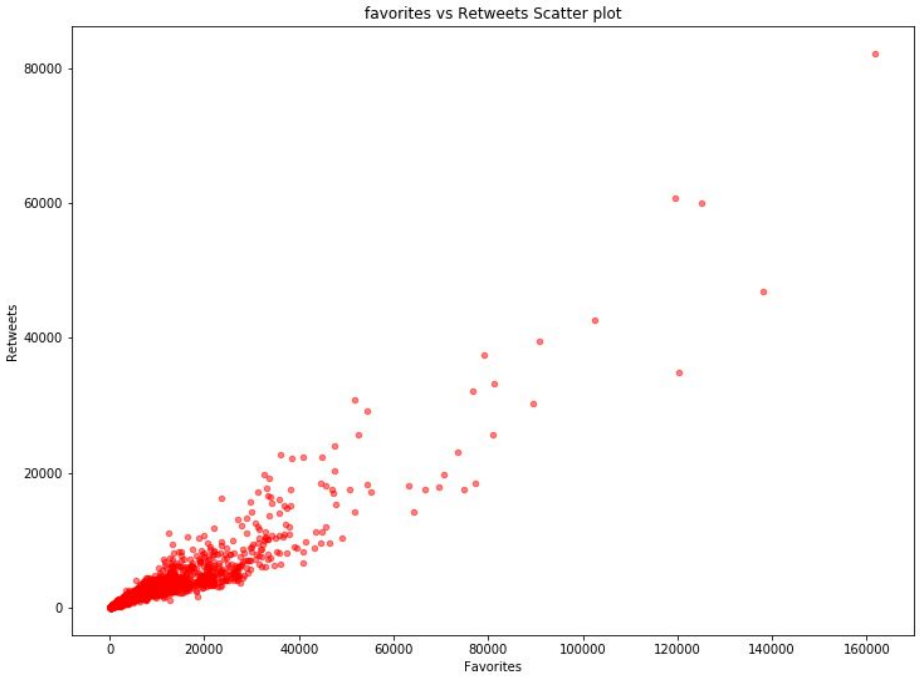
After tidying up the DataFrames i proceeded with data analysis and data visualisation.

1. **How the tweets posts trended over time.**



Inference: In the initial months of the given time period frequency of tweets per month was very high as compared to August 2017

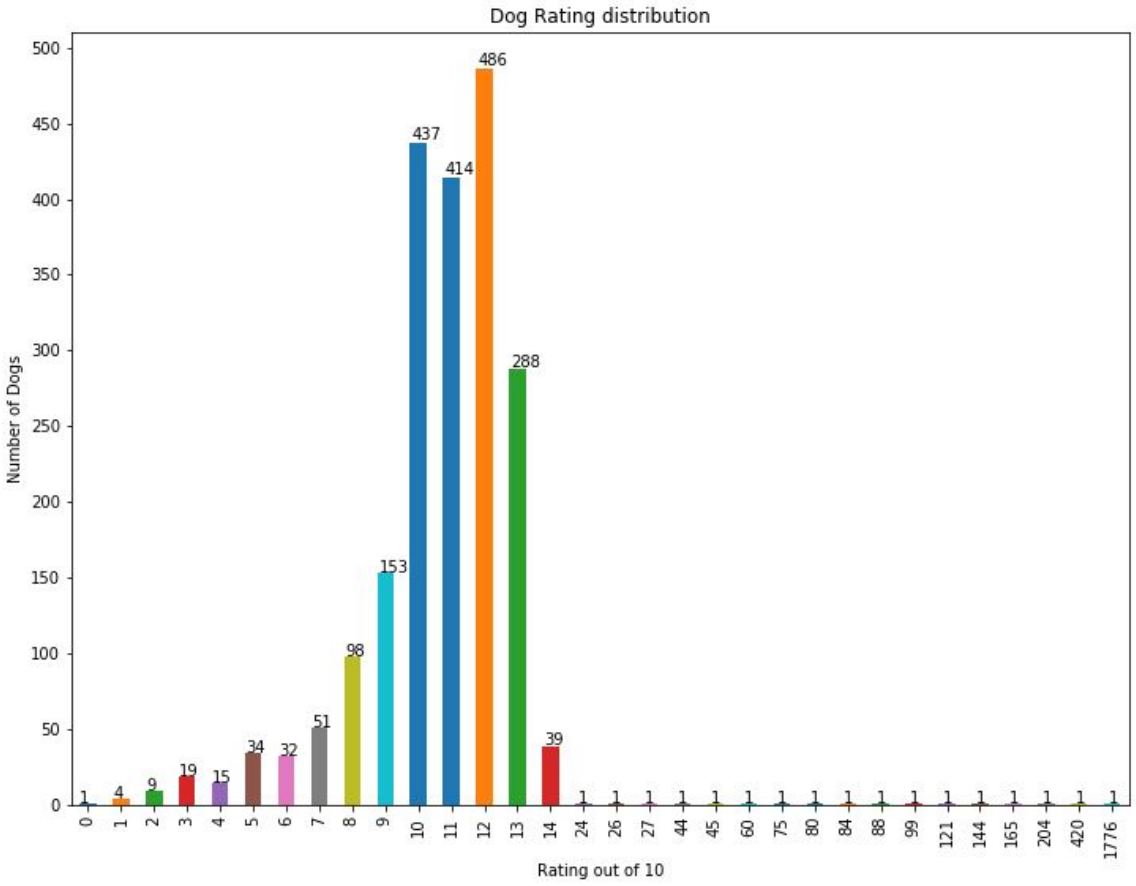
2 . Relationship between Retweets and favorites



Inference: From the above chart it is clear that there is a direct correlation between retweets and favorites.

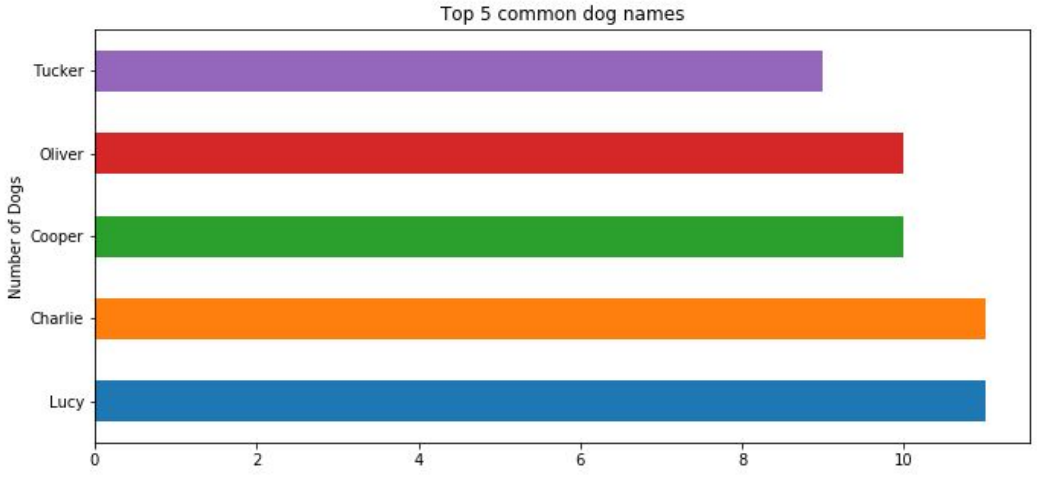
3. **Maximum retweeted tweet.**
Inference: The maximum number of retweet is: 82211.0, for the tweet: 825
https://twitter.com/dog_rates/status/744234799360020481/video/1

4**. Dog Numerator Ratings distribution**



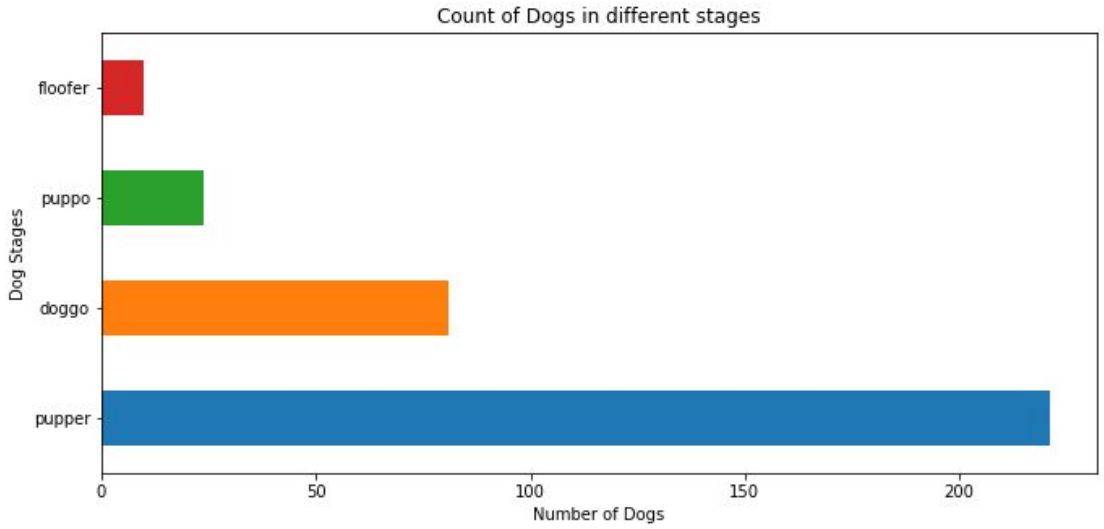
Inference: Numerator rating of 12 has been the maximum followed by rating of 10 and 11.

5. Top 5 common dog names



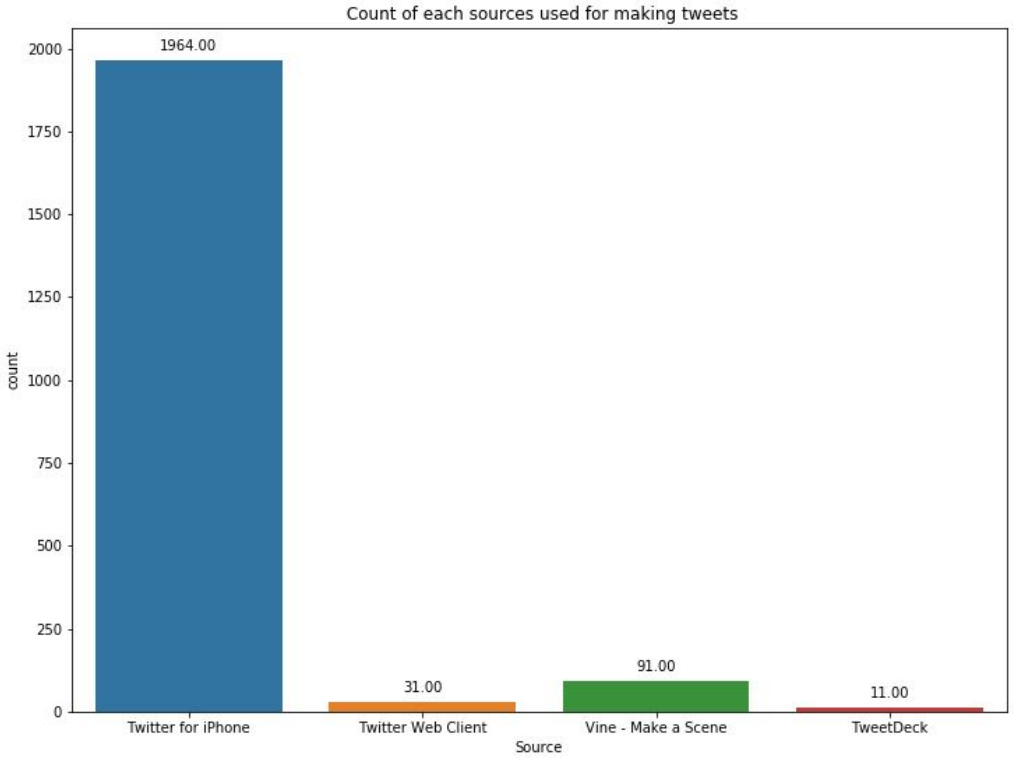
Inference: Lucy and Charlie are the most common Dog name.

6. Count of Dogs in different stages



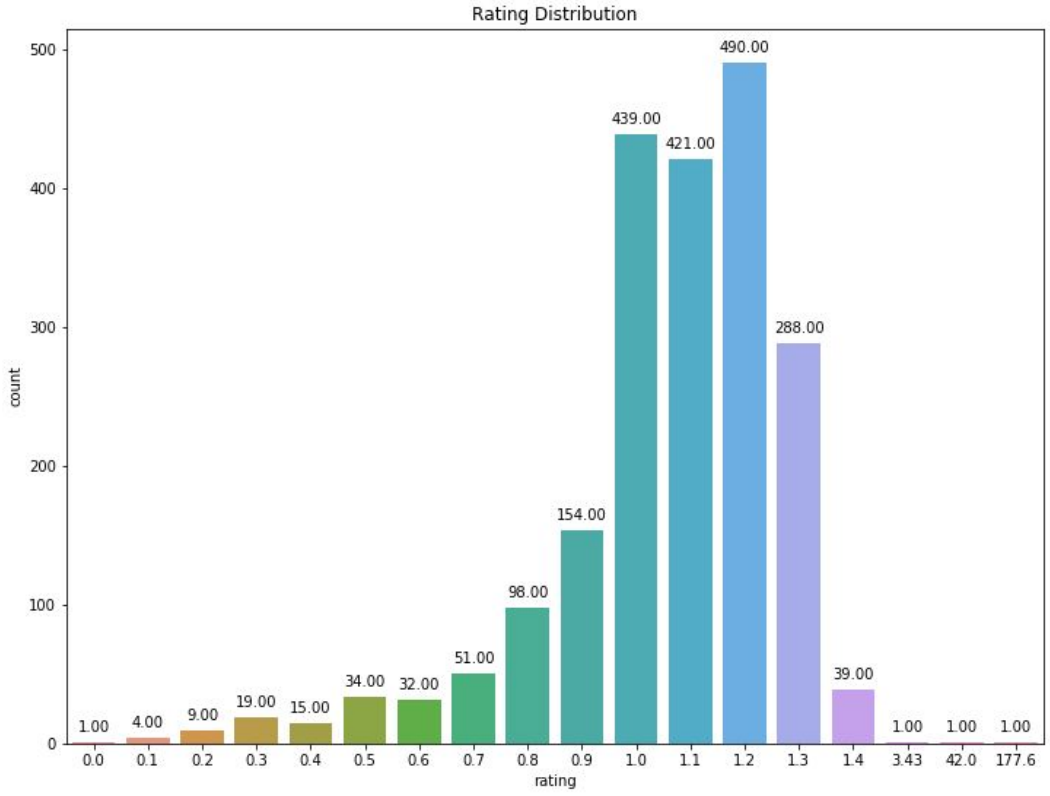
Inference: It seems most of the dogs in this data set are in Pupper stage.

7. Identifying the most used source for the tweets



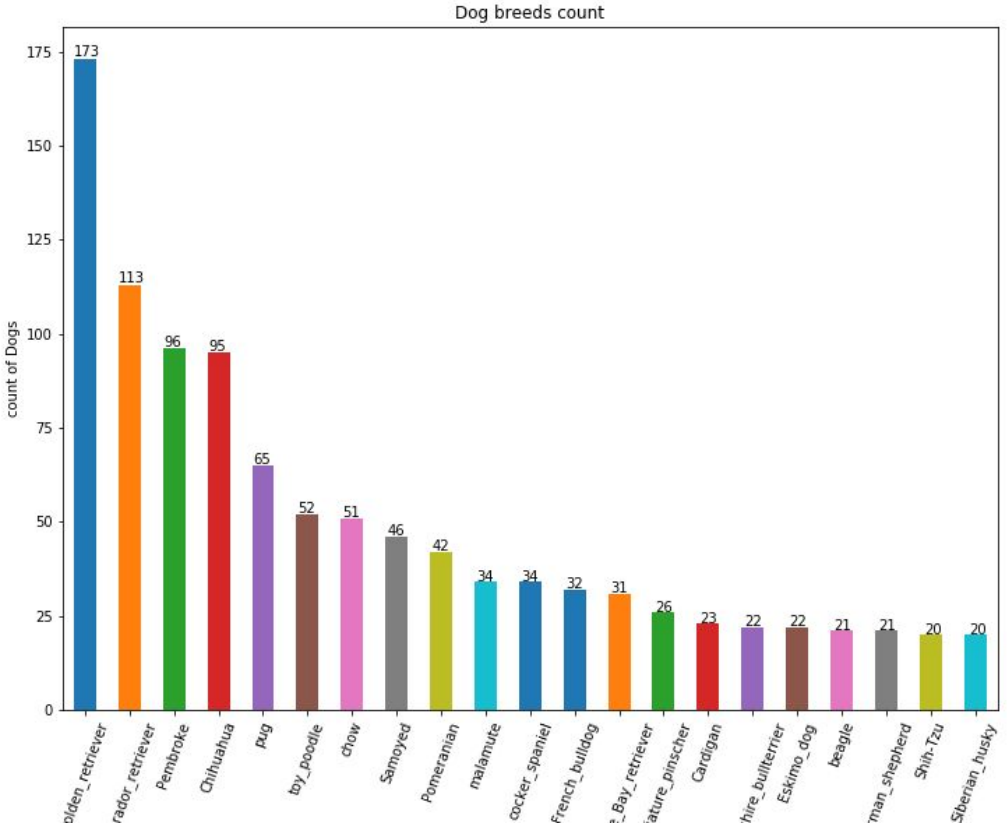
Inference: Most of the tweets have been made using iphone.

8. Calculated Rating distributions



Inference: Rating of 1.2 has been achieved by a large number of tweets followed by rating of 1.0 and then 1.1

9. Dog breed counts



Inference: Most of the dogs that are there in the dataset has been predicted to be Golden Retriever

10. Mean ratings

Inference: 1.165917978063899