Act report - We Rate Dogs Project

WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. The account was started in 2015.
- The tweeter archive enhanced file contains data extracted from 2356 of the 5000+ tweets account. The archive contains basic tweet data, for example dog name, dog 'stage' and dog raiting.

- From the file json data we extract retweets and favorites count.
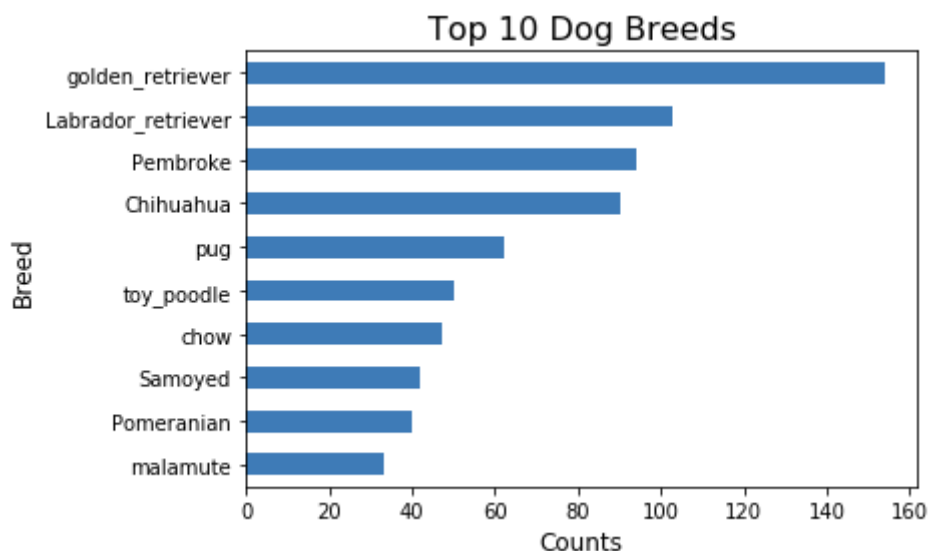
- The imagine prediction file gave us information about what brend of dog is present in each tweet. In particular it contains the top three brend predictions for each tweet ID, the image URL, and the image number corresponded to the most confident prediction.

Those dataset have been cleaned and merged in the dataframe tweeter_archive_master

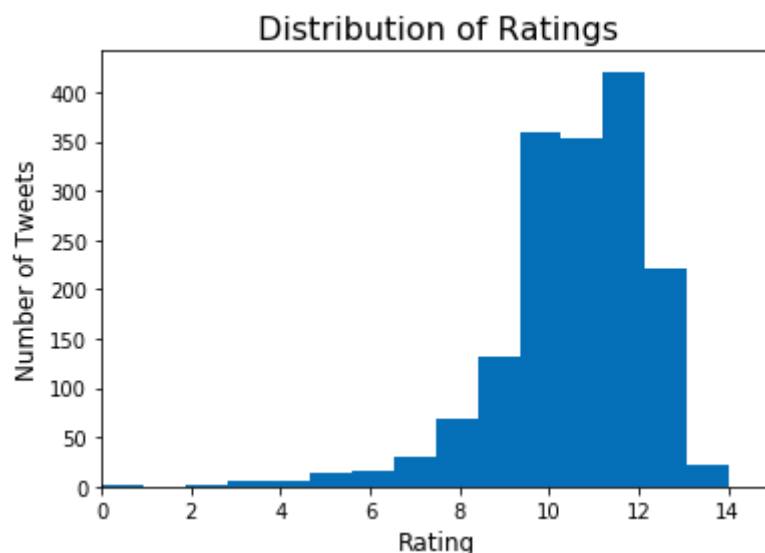In the section Exploratory Data Analysis we answered to the following questions:

1) Which breed are most popular?

The most common dog breed predected is the Golden Retriever with 154 tweets.
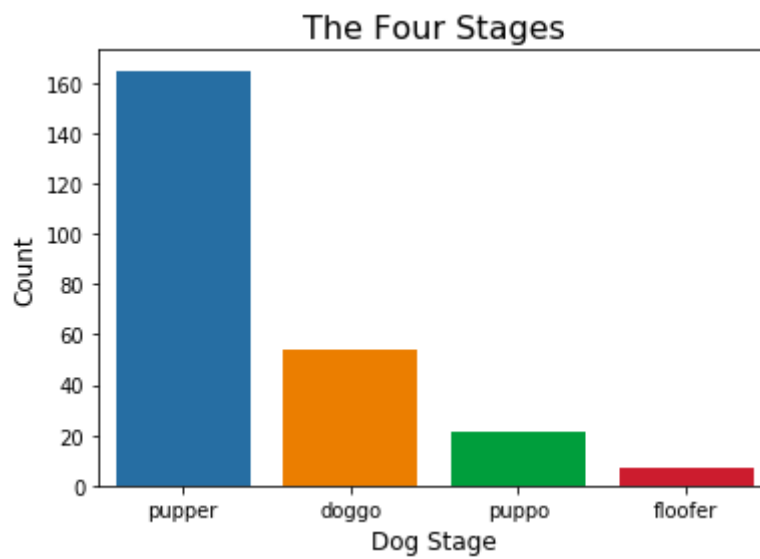

2) Which is the rating distribution?
We saw from the histogram and from the statistics that the 75% of the raiting is at 12.
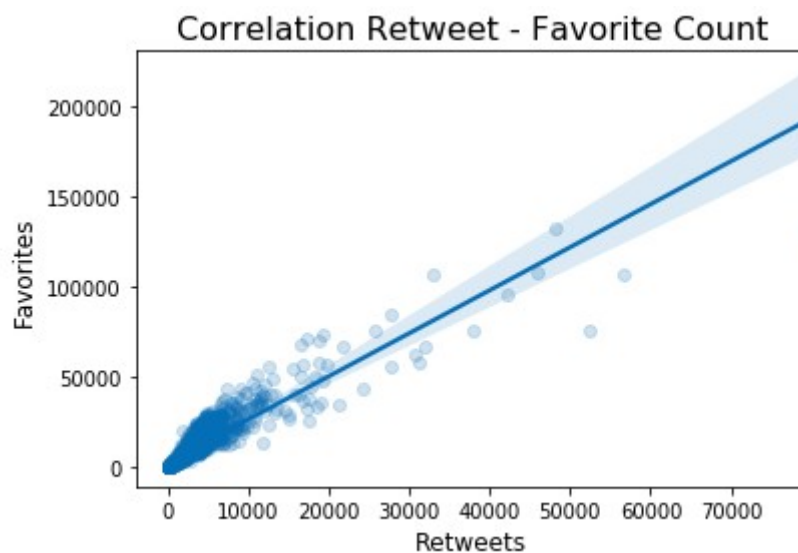
3) Which is the common stage?
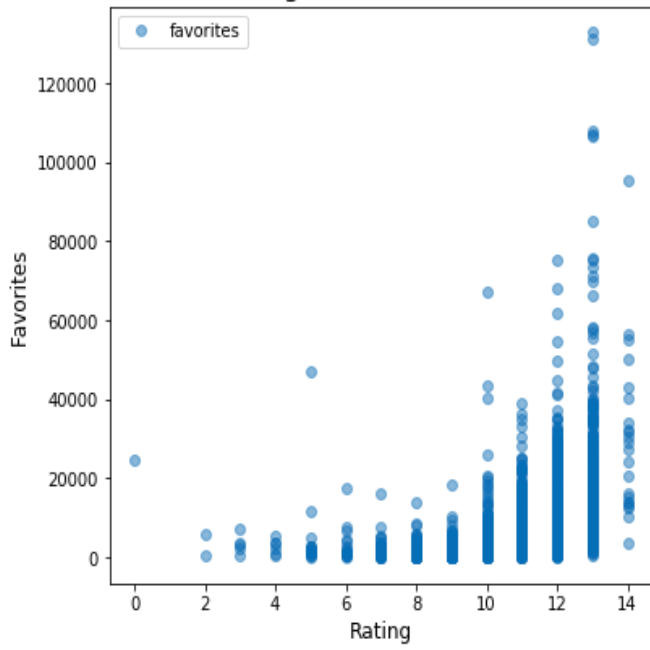Most dogs are classified in the 'Pupper' stage.

The Four Stages

4) Analysis: Correlation between retweets and favorites counts

It is obvious that when a picture is retweeted, it will be more favorite in contrast to other posts that receiced less rating and minus visibility. We can see those aspects the following grafics:

- 'Correlation between retweets and favorites posts': if the post is retweeted more times, it will get more visibility.
- 'Retwees counts - Ratings' shows that the haigher rated dogs receive more retweets.
- 'Favorites counts - Ratings' shows that the haigher rated dogs are more favorites.

Correlation Retweet - Favorite Count

Ratings v. Favorite Count

Ratings v. Retweet Count