Project 5: Udacity « Data Analyst »

The project was aimed at wrangling data collected from the WeRateDogs twitte database which is a Twitter account that rates people's dogs with a humorous commen about the dog. The data was cleaned to only 1366 unique tweet id with the rating, the stage and the predicted class of dogs as well with the number of retweets and favorite

The first question was who is the first dog in terms of rating, retweet and favorite: and it is Bo who got 14/10, 38488 retweets and 90190 favorites. The second one is Puppo and the third one Gabe (see tweet captures).
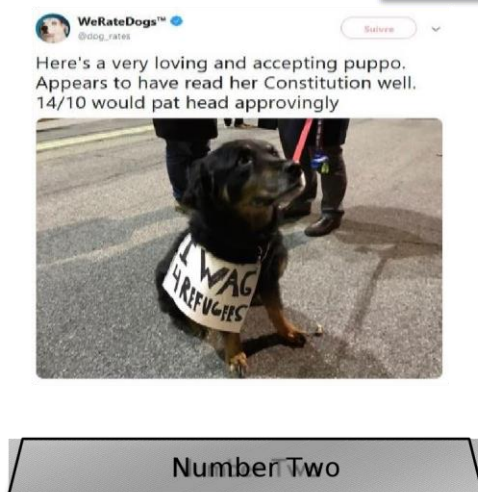
The next question is about the distribution of the rating: we have excluded the fantasists rating under 10 and above 20. There are five notes distributed differently: only few dogs got the 14/10, the note 12 was the one given the most.
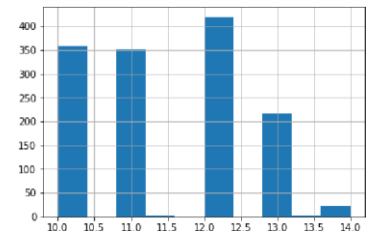
The question of the relation between rating, retweets was also investigated and appears that the most rated dogs are neither the most retweeted nor the most favorite Those who get the 13 grades were more known.
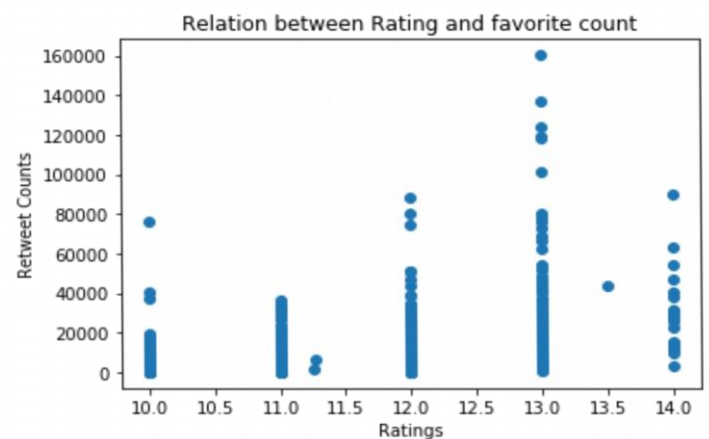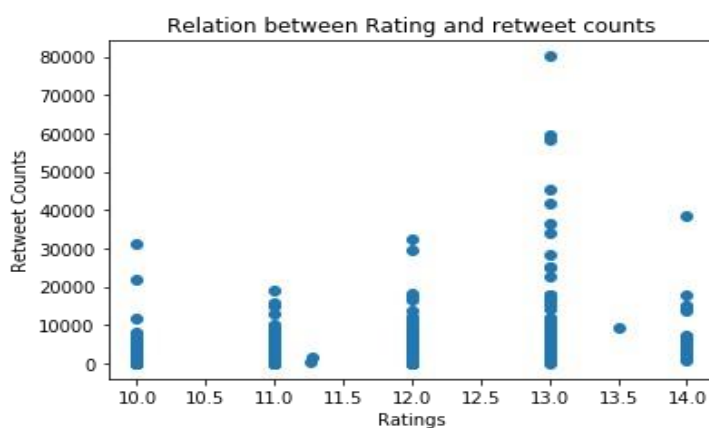
However, in average, the dogs who are more rated get more retweets and favorites.
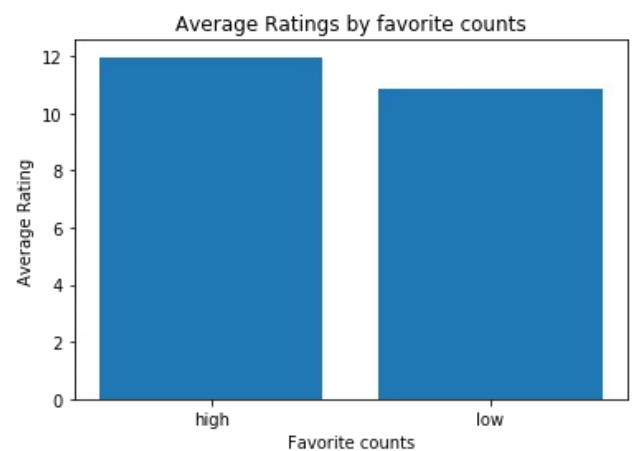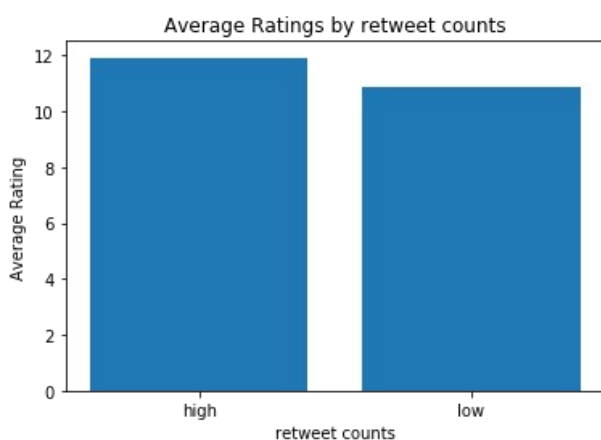
For dogs, like for humans, retweets and favorites are not the tools to rate even if they could be an indication.