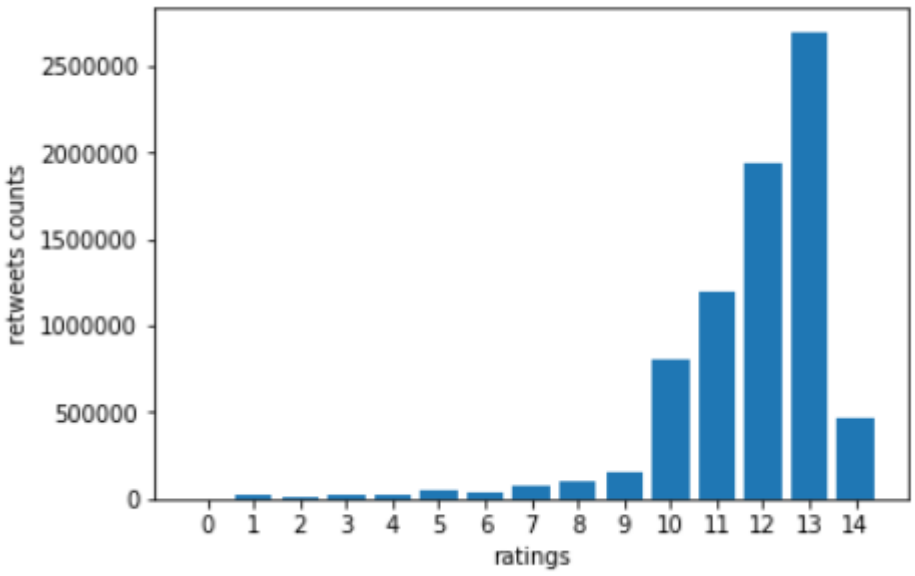
1. Analysis

After we finished cleaning the data, we need to start analyze it and to find insights into the data.

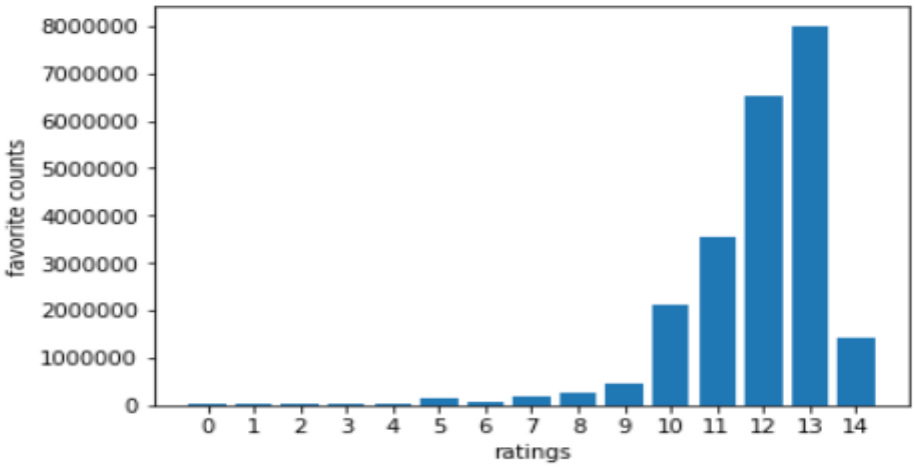
We need also to create visualization that would show hidden secrets in the data, we will go through steps to show how much work we have put into that project.

First step is to uncover data frames and review it in order to analyze it, we discovered the data, then after using some methods such as info() , describe , value counts method, we have reached some insight the numerator rating 12 is the mode of the data frame numerator ratings and 1776 is the highest rating was observed. I was curious about the neural network algorithms, so I challenged these algorithms with 1776 numerator rating, I used loc to find the row then I reached to the algorithms result the first algorithms was a bow tie and the second algorithms was sun glasses the third algorithms was sun glasses too. Then I visualized the photo it appeared that a dog was actually wearing a tie and sun glasses, the algorithms didn't catch the dog, so the third insight the algorithms aren't so accurate.

I did a visualization to relate the rating with retweet count, it showed that 13 is the highest retweet counts



I did another visualization to relate the rating with favorite count, it also showed that is the highest favorite counts is 13



I used describe method it showed that the algorithms one has the highest configuration and the mean of favorite count is higher than the mean of the retweet count. The highest mean retweeted messages of the algorithms the first thirty elements of all algorithms most of them are not actually a dog breed, for example the highest retweeted was an Arabian camel. The algorithms mayn't be very accurate,

I also relate dog type with retweet counts and favorite counts it showed the highest pair was (None,None)then(pupper,None)then(None,puppo) then(doggo,pupper) .