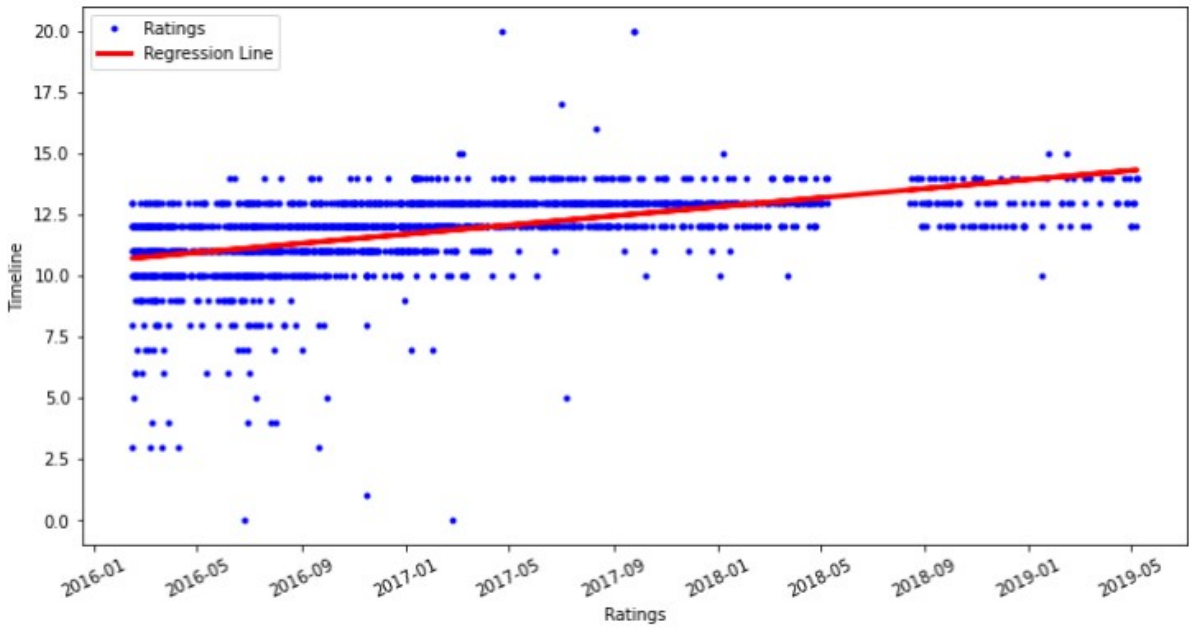
The graphing of the dog ratings is essentially a derivation of David H. Montgomery's Pup Inflation Post. He suggests that over the years, dog ratings on the internet has been getting better and better. To find out whether this is true we worked on a twitter data feed of dog ratings extracting the given ratings from the @dog_rates feed and performing Linear Regression on the data to find a trend between the predicted vs actual data.

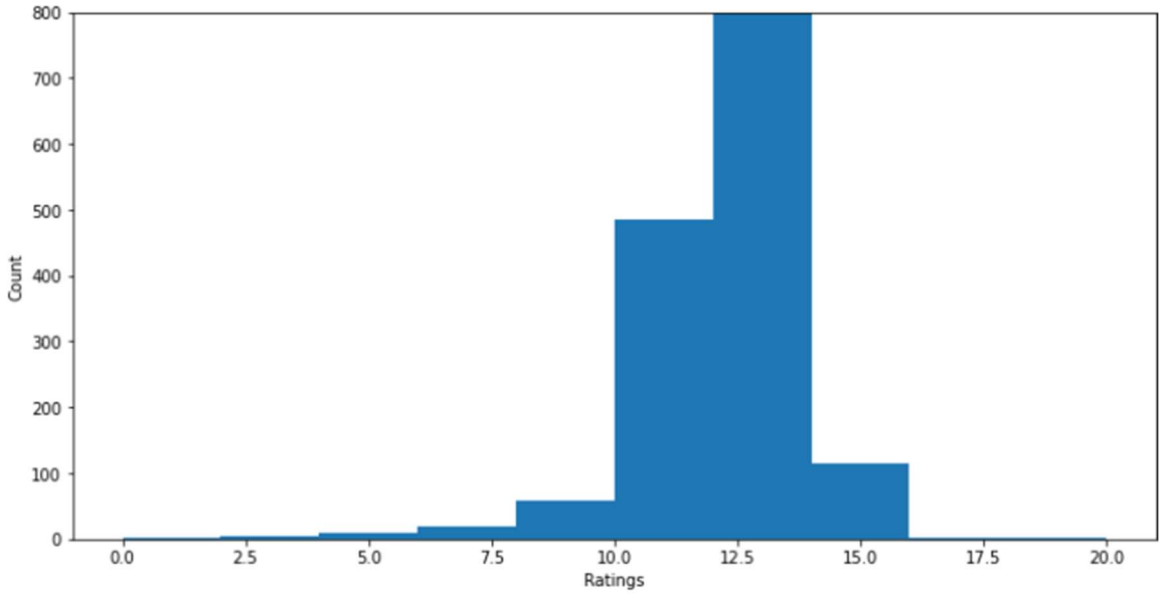
To create a usable dataset we first extract the ratings that actually make sense and removing any outrageous datapoints to avoid skewing the actual information we want to gather, so we keep ratings in a certain range <25 and remove the outliers. Apart from this, we extract the datetime to make it iteratable with the data so that it's simpler to put into the model and generate predictions. Now we put this into a linear regression model to calculate slope and intercept predictions. On the basis of these prediction we plot a linear regression line on the data points and the graph gives a clear picture of the relationship between the actual values vs predicted data.

The histogram on the other hand tells where the actual density of ratings lie which is around 12.5, it also goes on to show most people like dogs and some really really like them but it is a gradual trend on the lower ratings so less people hate dogs than the people who like them.

Graph 1



Graph 2


800
700
600
500
400
300
200
100
0
Count
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Ratings