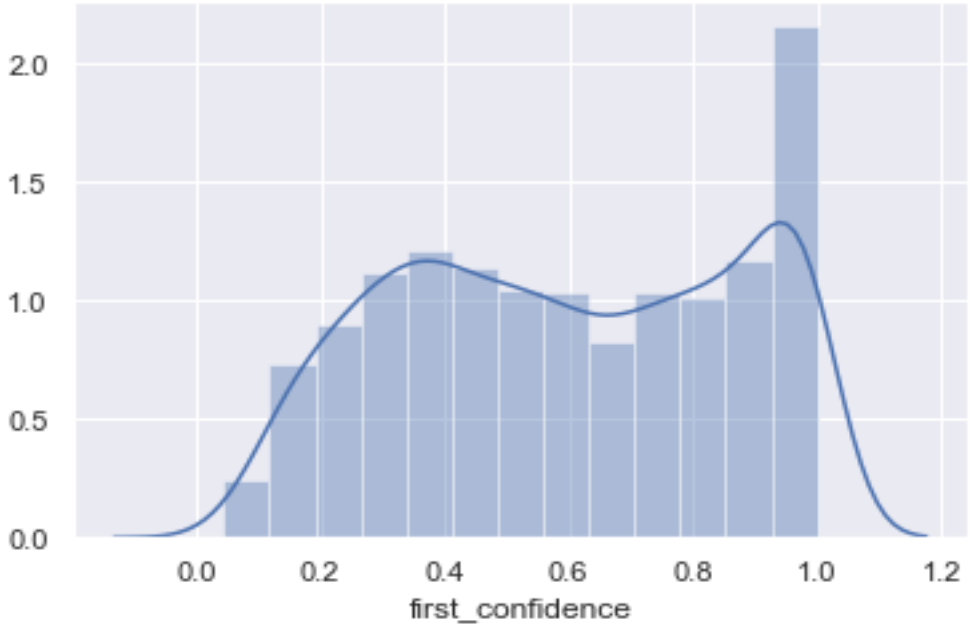
Data Analysis and Visualization

This report contain some insights about data analysis of Udacity project We-Rate-Dogs and it contains three insights and visualizations

1- The Confidence of the model

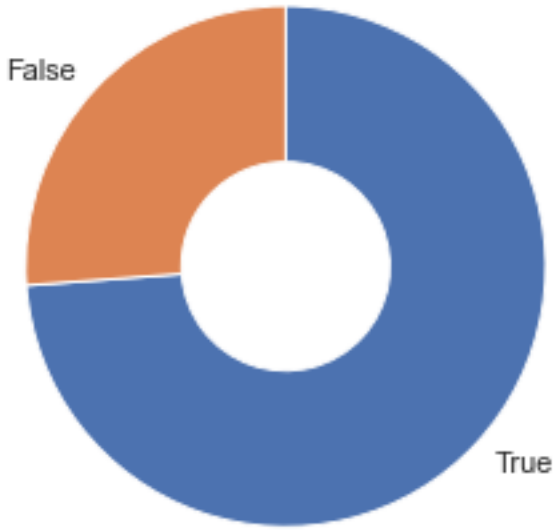


The above figure shows the confidence of first prediction of the algorithm i.e. whether the system is confident about prediction of a kind of a dog or not. It is very clear that we have many cases of 100% but the wide range from 0.2 to 0.8 is telling us about the low quality of our system prediction.

golden_retriever	150
Labrador_retriever	100
Pembroke	89
Chihuahua	83
pug	57
chow	44
Samoyed	43
toy_poodle	39
Pomeranian	38
cocker_spaniel	30
malamute	30
French_bulldog	26
miniature_pinscher	23
Chesapeake_Bay_retriever	23
seat_belt	22
Staffordshire_bullterrier	20
Siberian_husky	20
German_shepherd	20
Cardigan	19
web_site	19

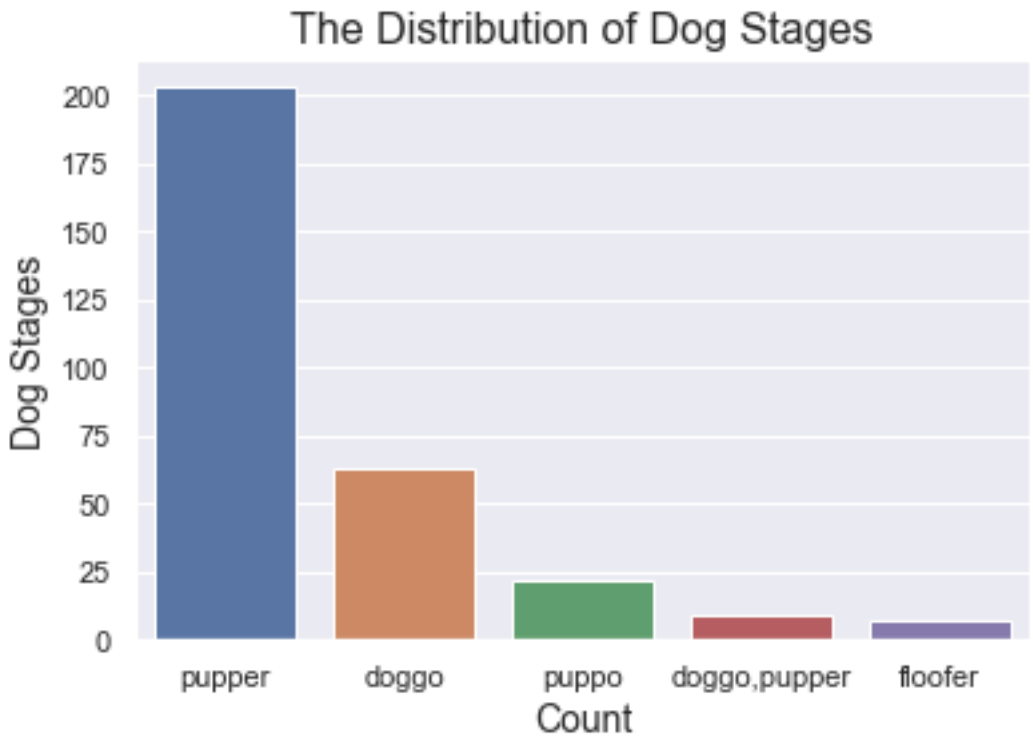
This model prediction is based on the above breeds of dogs which are the most popular breeds in our data. Golden retriever and Labrador are the top two breeds as it is very clear from our data analysis and both of them have at least 100 predictions.

2- The first prediction of the model



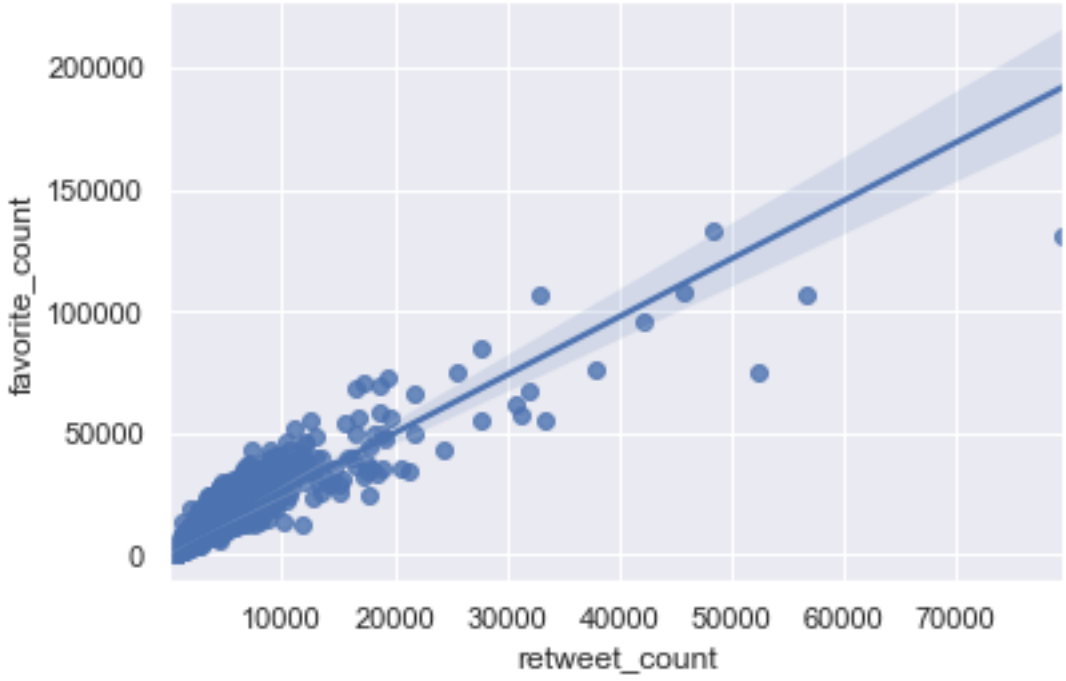
The pie chart is conveying that around 70% of the first prediction of being a type of a dog or not is correct. This success rate in prediction is an indication that our system needs to be enhanced to a quality that is suitable for a deep learning model.

3- The distribution of the dog stages



From the bar chart of the distribution of dogs it is very obvious that the pupper (small dog) is the most popular breed between all types followed by the doggo, puppo, doggopupper and floofer. This means that the younger the dog, the more popular it is. We can also say that our model is not reflecting the truth because we have many missing stages in this distribution

4- Favorite count and retweet count relationship



The above figure implies that favorite posts have more retweets and we can say that the more the increase in favorite counts, the more the increase in retweets.