

# Act report

## INTERDICTION :

After going through all of the data wrangling steps on the dataset which is the tweet archive of Twitter user @dog\_rates, also known as WeRateDogs. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog, i can perform some data analysis and visualizations which i will demonstrate below.

### 1. insights

at first I looked at what year the account had a high retweets amounts after calculating the sum of retweets for each year the fig(1) shows that the amount of retweets for tweets created in 2016 is the highest , which I am assuming its because the account was more active or the data was biased.

years	
2015	717925
2016	3004222
2017	2447286

fig(1)

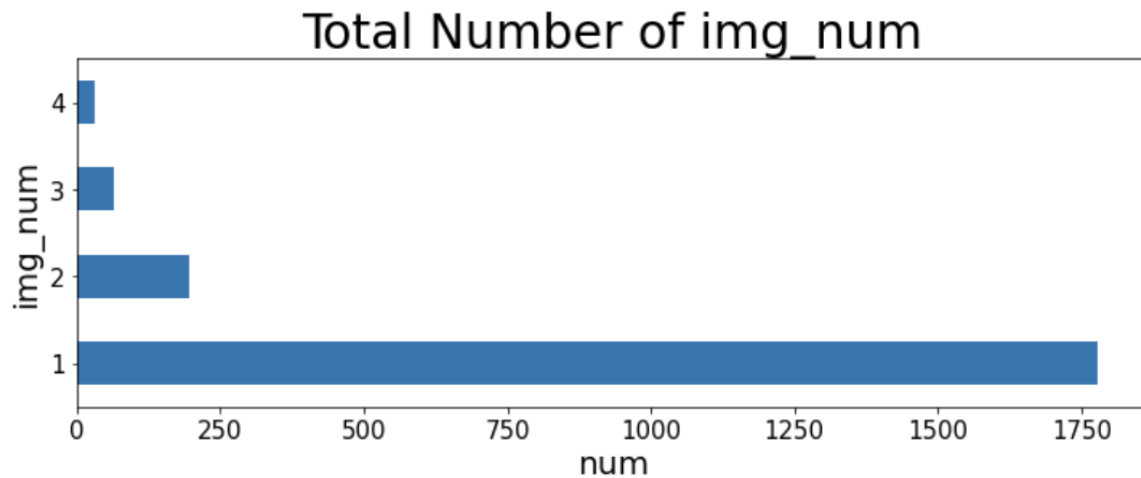
I also was very interested to know which algorithm has the highest prediction confidence rate, so I calculated the mean for each algorithm prediction confidence and it shows that the first algorithms has the highest prediction confidence.

lang	
en	2065
et	1
eu	1
in	2
nl	3
ro	1

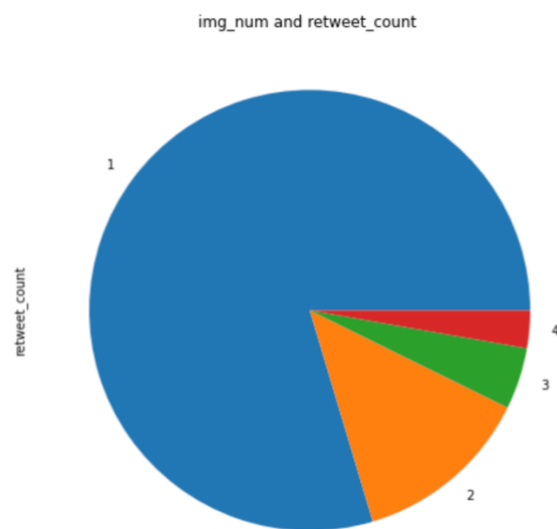
fig(2)

after counting how many tweets in each language the fig(2) shows that the majority of tweets are written in English.

### 2.visualization



according to our findings in fig(3) users are more likely to share one photo of their dogs,



Fig(4)

I was also curious in how the amount of images uploaded by users could affect the number of retweets.

and it shows In fig(4) the number of retweets for tweets that feature one photo has the more numbers of retweets

I was initially assuming that having more images would assist with obtaining more retweets, but I suppose I was mistaken.