EXTERNAL DOCUMENT

Act Report

By

Mohamed Sobhy

Here I would like to mention that there are many valuable insights and I chose the insights I included in my project jupyter notebook file and the visualization I provided at end of it.

1- I looked at favorite tweet count distribution

	count	mean	std	min	25%	50%	75%	max
1992	8687.33	12742.54	78	1862.75	3929	10824	163487	

2- Let's see the dog with max number of favorite tweet

dog_stage	favorite_count	dog_prediction	confidence
doggo	163487	Labrador Retriever	0.825333

3- I also looked at retweet count distribution

	count	mean	std	min	25%	50%	75%	max
1992	2634.65	4697.23	11	588	1265	3019	83127	

4- And guess what!, the dog with max favorite count is the same with max retweet count

dog_stage	retweet_count	dog_prediction	confidence
doggo	83127	Labrador Retriever	0.825333

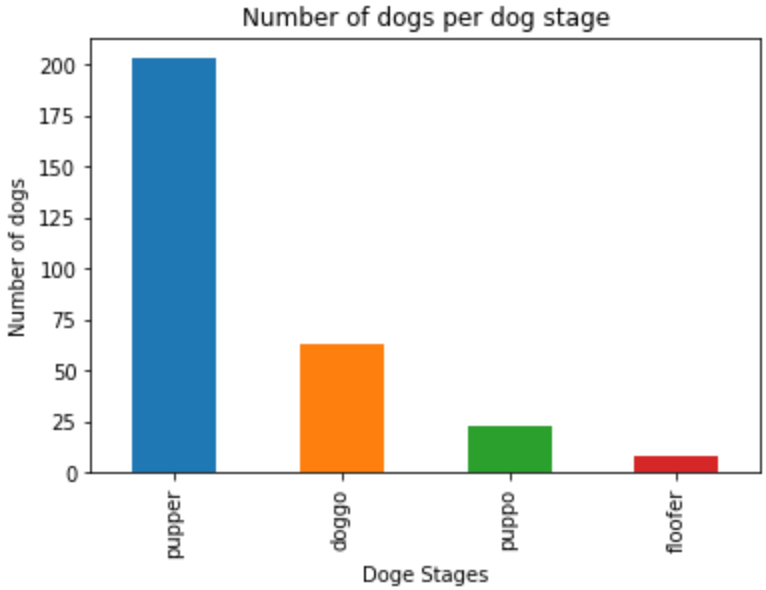
5- What is the distribution of prediction confidence level

	count	mean	std	min	25%	50%	75%	max
1992	0.465	0.339	0.000	0.140	0.458	0.776	0.999	

6- The dog that has the max confidence level is the d=following:

source	favorite_count	retweet_count	dog_prediction	confidence
Twitter for iPhone	1051	497	Komondor	0.99996

7- I put a visualization that shows the number of dogs per each dog stage after excluding the unidentified dog stage “None”. It shows that **pupper** is the most dog stage between the ranked dogs



8- Due to lack of time before submission I wish I had the chance to provide more insights and visualization which I only could show you some of them on the following tableau url.

https://public.tableau.com/profile/mohamed.sobhy#!/vizhome/WeRateDogs_15515507995680/Dashboard1?publish=yes