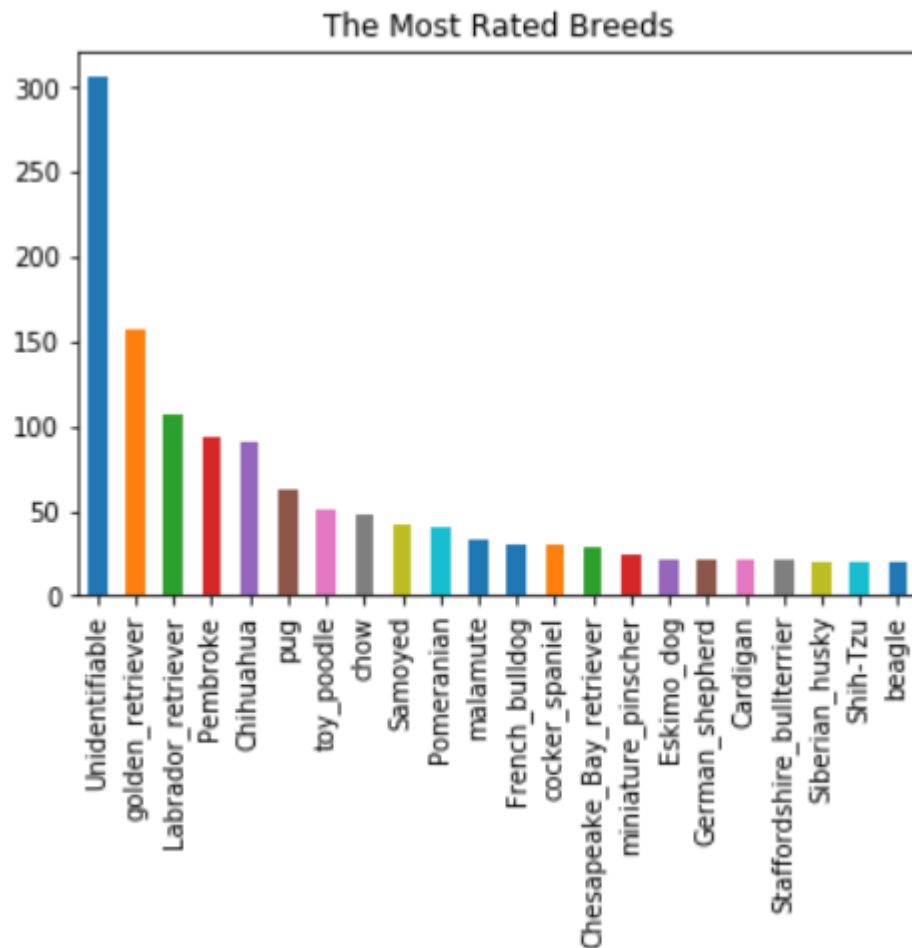


Act Report

We have taken the twitter data from twitter_archive_master.csv which has 1982 entries. We try to find the most rated Breeds



It's difficult to know which of these breeds are the top breeds. Because majority of them are unidentified we need to improve the model which is predicting them.

	count	mean	std	min	25%	50%	75%	max
breed								
Cardigan	21.0	1.114286	0.159015	0.7	1.000	1.10	1.2	1.3
Chesapeake_Bay_retriever	29.0	1.082759	0.146553	0.8	1.000	1.00	1.2	1.3
Chihuahua	91.0	1.050949	0.209656	0.3	1.000	1.10	1.2	1.4
Eskimo_dog	22.0	1.140909	0.140269	0.9	1.000	1.20	1.2	1.4
French_bulldog	30.0	1.116667	0.182101	0.8	1.000	1.15	1.2	1.4
German_shepherd	21.0	1.100000	0.144914	0.8	1.000	1.10	1.2	1.3
Labrador_retriever	107.0	1.119626	0.132066	0.7	1.000	1.10	1.2	1.4
Pembroke	93.0	1.141935	0.173407	0.4	1.100	1.20	1.2	1.4
Pomeranian	41.0	1.095122	0.139555	0.6	1.000	1.10	1.2	1.4
Samoyed	42.0	1.169048	0.135229	0.7	1.100	1.20	1.3	1.4
Shih-Tzu	20.0	1.035000	0.130888	0.8	0.975	1.05	1.1	1.3
Siberian_husky	20.0	1.125000	0.116416	0.9	1.000	1.10	1.2	1.3
Staffordshire_bullterrier	21.0	1.076190	0.137495	0.8	1.000	1.10	1.2	1.3
Unidentifiable	306.0	0.900327	0.326280	0.0	0.700	1.00	1.1	1.4
beagle	20.0	1.015000	0.153125	0.6	0.975	1.00	1.1	1.3
chow	48.0	1.141667	0.135007	0.7	1.100	1.20	1.2	1.3
cocker_spaniel	30.0	1.133333	0.118419	0.9	1.025	1.15	1.2	1.3
golden_retriever	157.0	1.152345	0.132511	0.5	1.100	1.20	1.2	1.4
malamute	33.0	1.087879	0.121854	0.8	1.000	1.10	1.2	1.3
miniature_pinscher	25.0	1.000000	0.256580	0.2	0.900	1.10	1.2	1.2
pug	62.0	1.024194	0.181691	0.3	0.925	1.00	1.1	1.3
toy_poodle	51.0	1.103922	0.126429	0.7	1.000	1.10	1.2	1.3

count 1982.000000

mean 1.053963

std 0.219475

min 0.000000

25% 1.000000

50% 1.100000

75% 1.200000

max 1.400000

Name: rating, dtype: float64

1.200000	452
1.000000	420
1.100000	399
1.300000	261
0.900000	151
0.800000	95
0.700000	51
1.400000	35
0.500000	34
0.600000	32
0.300000	19
0.400000	15
0.200000	10
0.100000	4
0.000000	2
0.636364	1
0.818182	1

Name: rating, dtype: int64

- Here we have a statistical comparison of the top breeds with all the ratings. Only the top breeds have a mean higher than the total population mean.
- Here mean rating is around 1.053963. The 3 most common ratings are 1.2,1.0,1.1

Among the 4 dog stages, pupper has the biggest frequency, but pupper also gets the lowest favorite counts and retweet counts and rating.

Here are the visualisationsto support our claim.

