ACT_REPORT

Insights;

1. From the dataset it can be deduced that Pupper is the most common dog stage followed by doggo, puppo and lastly floofer. It should however be noted that there were many missing values and therefore the pupper may not accurately be the most popular. The pupper is the youngest do stage so in this scenario it can be presumed that most people prefer puppers as they are small and adorable compared to the other stages.

Pupper	209
Doggo	62
Puppo	23
Floofer	8

2. The most common dog names are Charlie, Cooper and Oliver. Also to note is that there were many null values for dog names meaning that the data could be misleading

Charlie	10
Cooper	10
Oliver	10
Lucy	9
Tucker	9
Pennie	9
Saddie	8
Winston	8
Lola	7
Daisy	7

3. The correlation between retweet count and favorite count is 0.93 which is very high. This means there is a strong relationship between the 2

	img_num	p1_conf	p1_dog	p2_conf	p2_dog	p3_conf	p3_dog	retweet_count	favorite_count	rating
img_num	1.000000	0.202714	0.032714	-0.156567	0.056999	-0.137806	0.059547	0.105033	0.130842	-0.000179
p1_conf	0.202714	1.000000	0.127061	-0.511022	0.142213	-0.707314	0.118888	0.047262	0.069590	-0.008762
p1_dog	0.032714	0.127061	1.000000	0.110839	0.640160	0.049916	0.560674	0.000884	0.042094	-0.029773
p2_conf	-0.156567	-0.511022	0.110839	1.000000	0.096856	0.481933	0.064919	-0.014106	-0.016257	-0.013995
p2_dog	0.056999	0.142213	0.640160	0.096856	1.000000	0.033872	0.559345	0.012614	0.052188	-0.035045
p3_conf	-0.137806	-0.707314	0.049916	0.481933	0.033872	1.000000	0.035461	-0.038351	-0.046251	-0.005131
p3_dog	0.059547	0.118888	0.560674	0.064919	0.559345	0.035461	1.000000	0.005500	0.033797	-0.030976
retweet_count	0.105033	0.047262	0.000884	-0.014106	0.012614	-0.038351	0.005500	1.000000	0.928502	0.022705
favorite_count	0.130842	0.069590	0.042094	-0.016257	0.052188	-0.046251	0.033797	0.928502	1.000000	0.021574
rating	-0.000179	-0.008762	-0.029773	-0.013995	-0.035045	-0.005131	-0.030976	0.022705	0.021574	1.000000

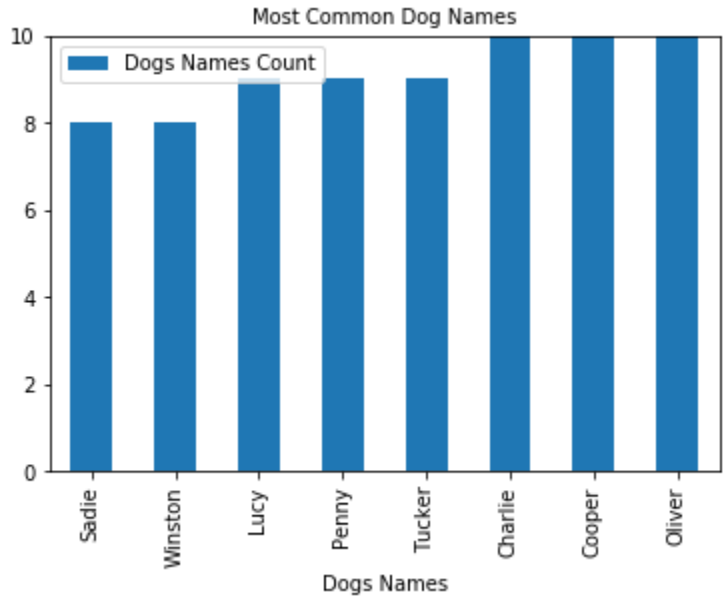
4. Dog ratings are averagely more than 1, which is understandable as most dog owners will give their dogs a high rating

count	1963.000000
mean	1.169375
std	4.098957
min	0.000000
25%	1.000000
50%	1.100000
75%	1.200000
max	177.600000

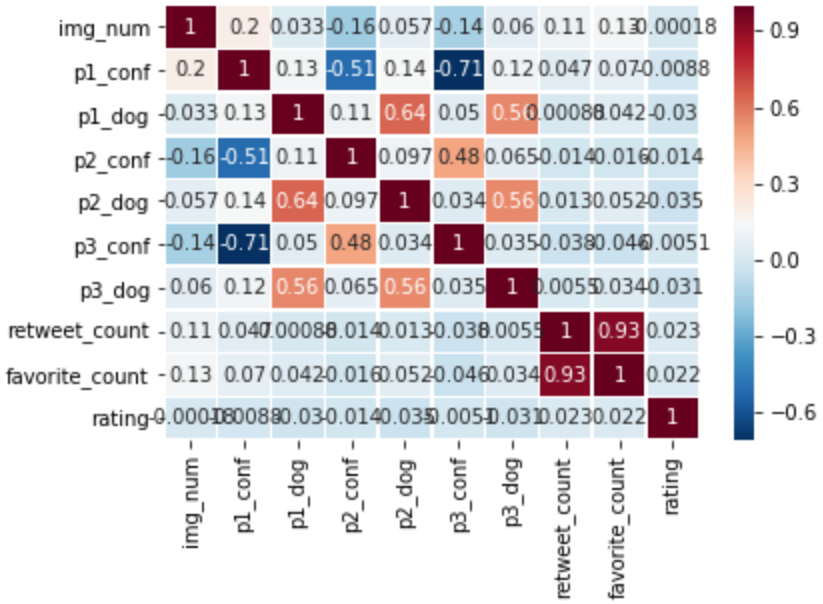
Visualization

Most common Dog Names Bar Graph

The top 10 most common dog names have been displayed as per below

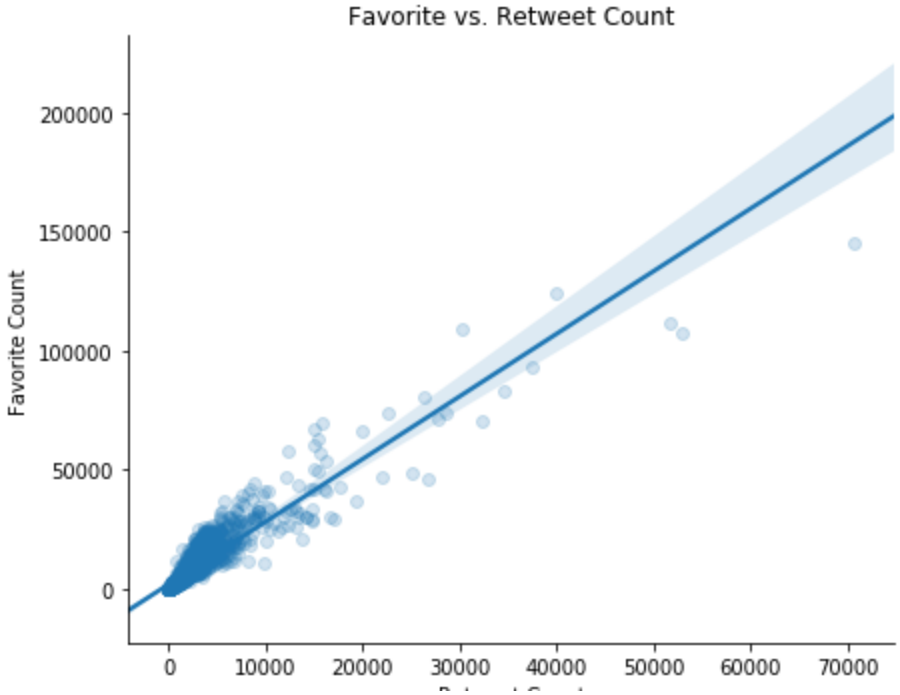


The correlation coefficients can be visualized here between different numeric variables. Comparing retweet count and favorite count, the correlation coefficient is 0.93 (close to 1) and we can therefore deduce that there is a strong correlation between the 2



Favorite vs Retweet Count Line Graph

There is a strong positive linear relationship between favorite count and retweet count. This means tweets that are liked (favorite) are likely to be retweeted. The coefficient for this is 0.93, showing that it is indeed true


Favorite vs. Retweet Count
Favorite Count
200000
150000
100000
50000
0
0
10000
20000
30000
40000
50000
60000
70000