Wrangle and Analyse Data Insights

Act report DAND Project

Did you ever wonder what a good twitter advertisement strategy for a dog product might look like? In this report I argue for a data driven decision approach in this matter, based on the output of the DAND 4th project analysis phase.

The Design of your dog persona:

A golden retriever pupper named Charlie seems to be the perfect dog for resonating the most with the twitter dog lovers community.

Counting the used names in the dataset results in a clear favourite “Charlie” followed by “Oliver” and “Cooper”. As you can see in *Figure 1 left.* Sorting the stage of development by amount of tweets in *Figure 1 right* most tweets are about “puppers” with overwhelming 66.8%. In opposite ‘floofers’ are getting the least attention on twitter with 2.5%. Finally, the most representative breed (by counting the breed of all the tweets with a rating higher than 12) is the golden retriever.

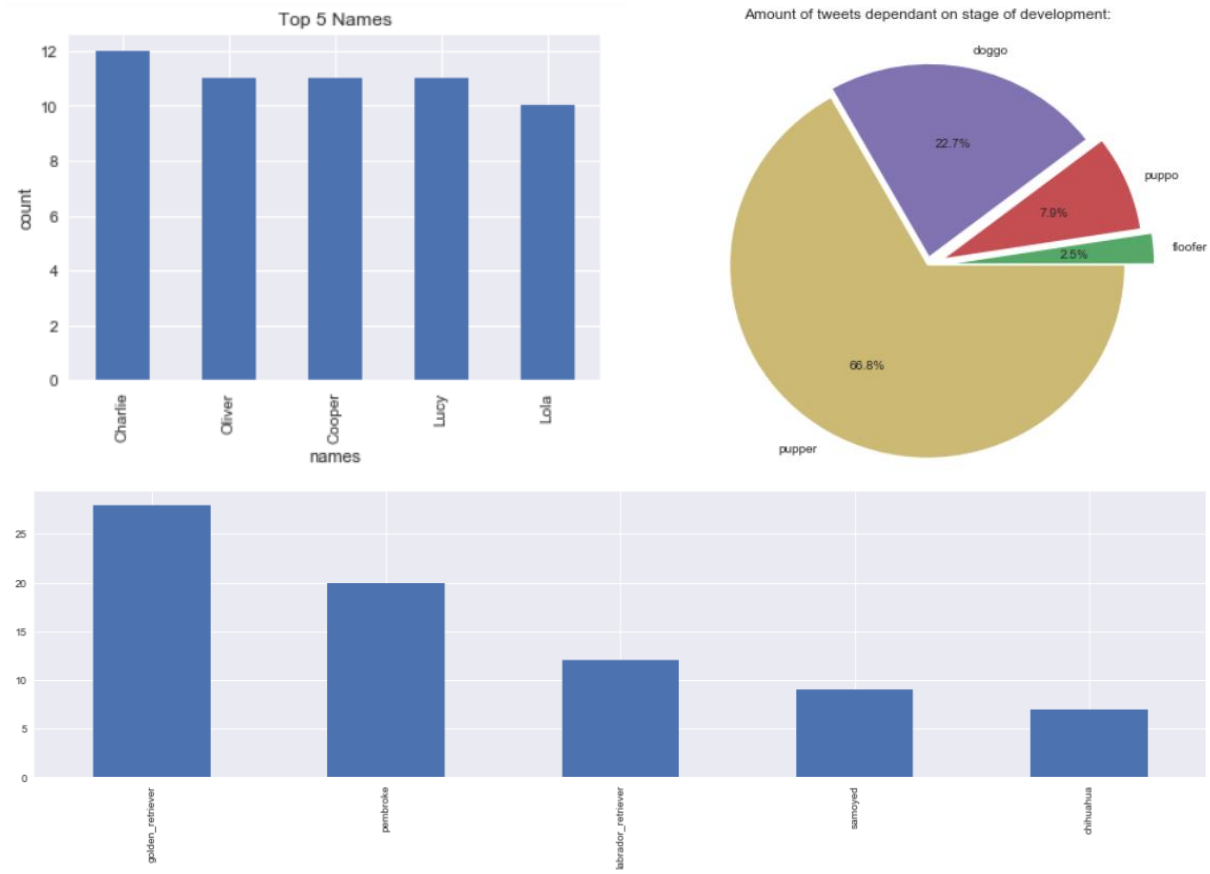

*Figure 1: Top 5 names (**left**), tweet amount in relation to development stage (**right**), top 5 breeds with highest ratings (**below**)*

The strategy design:

Post your post on a Sunday from 18 - 24 o'clock to avoid high post traffic and expect the most visibility on Wednesday.

To increase the chance of attention it's better to post on a day with less tweets. As shown in Figure 2 (**right**) the least posts are made on Sunday. Most tweets are tweeted between 0 and 6 o'clock, Figure 2 (**below**). After 6 and before 12 o'clock there only a few tweets. Which shows that it is the worst time to attract attention on this dog channel. A moderate amount of tweets with related high ratings can be found after 18 0'clock. And should you have a highly rated tweet and you want to reach many people then you should wait until Wednesday since the correlation for retweet count and rating is the highest on this day as can be seen in Figure 2 (**right**).

weekday		retweet_count	rating_numerator
Fri	retweet_count	1.000000	0.216878
	rating_numerator	0.216878	1.000000
Mon	retweet_count	1.000000	0.290271
	rating_numerator	0.290271	1.000000
Sat	retweet_count	1.000000	0.223220
	rating_numerator	0.223220	1.000000
Sun	retweet_count	1.000000	0.305366
	rating_numerator	0.305366	1.000000
Thu	retweet_count	1.000000	0.360970
	rating_numerator	0.360970	1.000000
Tue	retweet_count	1.000000	0.369091
	rating_numerator	0.369091	1.000000
Wed	retweet_count	1.000000	0.380966
	rating_numerator	0.380966	1.000000

weekday	count
Fri	287.000000
Mon	333.000000
Sat	258.000000
Sun	254.000000
Thu	288.000000
Tue	303.000000
Wed	311.000000

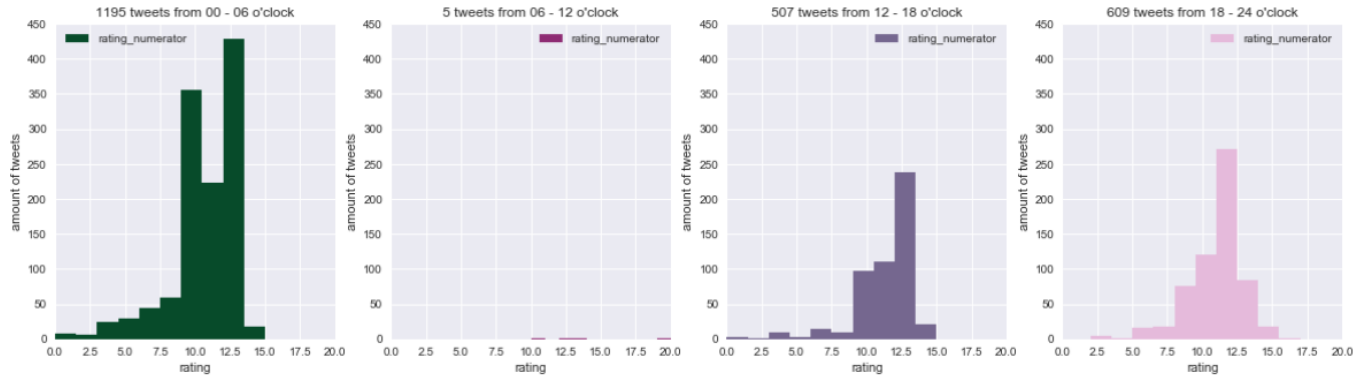

*Figure 2: Correlation between rating and retweet count ordered by weekday (**left**), amount of tweets by weekday (**right**), amount of tweets vs ratings by time of the day (**below**)*