

# Analyzing And Visualizing WeRateDogs.

*By Lucy Nsan*

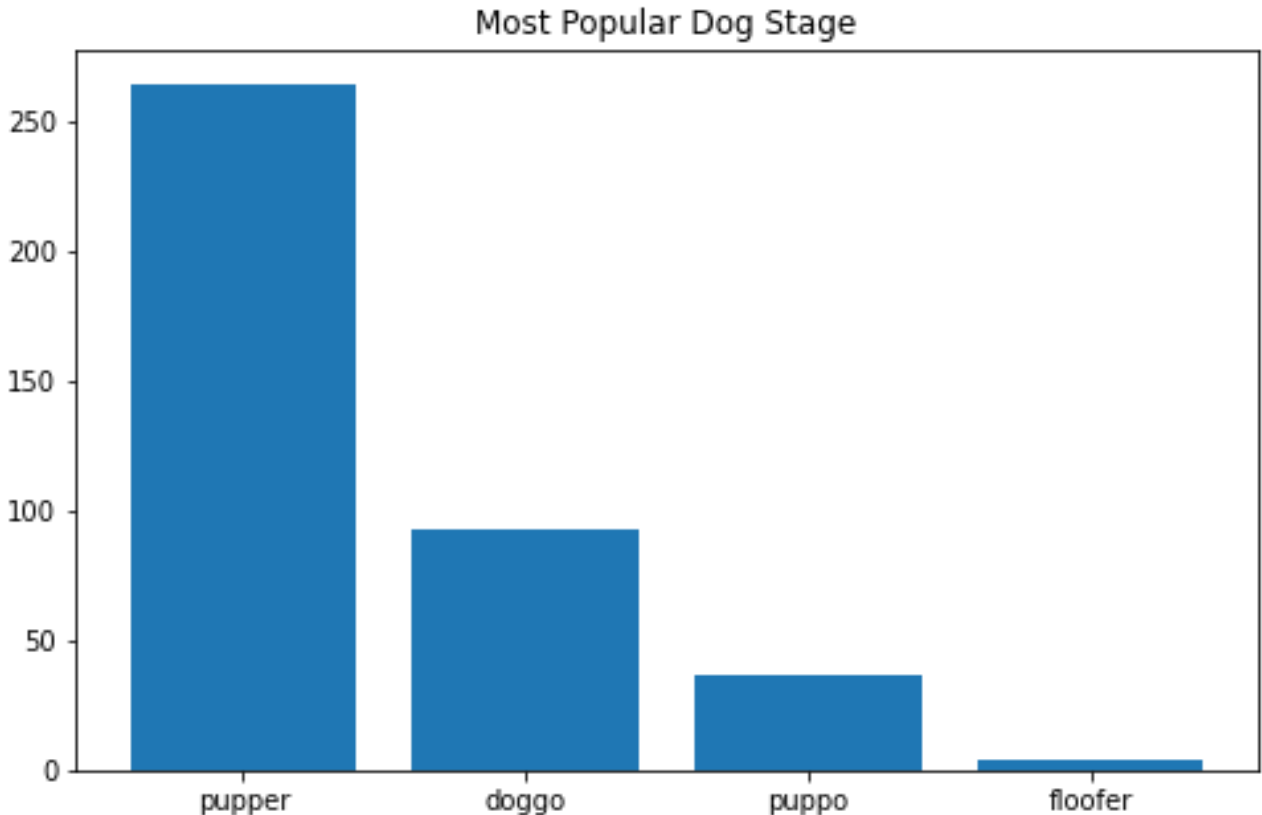
## Introduction

WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings always have a denominator of 10. The numerators though, can be greater than 10. 11/10, 12/10, 13/10, etc. This is part of the humour that makes WeRateDogs popular with over 6 million followers and has received international media coverage.

Udacity downloaded their Twitter archive exclusively to use in this project. This archive contains basic tweet data (tweet ID, timestamp, text, etc.) for all 5000+ of their tweets as they stood on August 1, 2017.

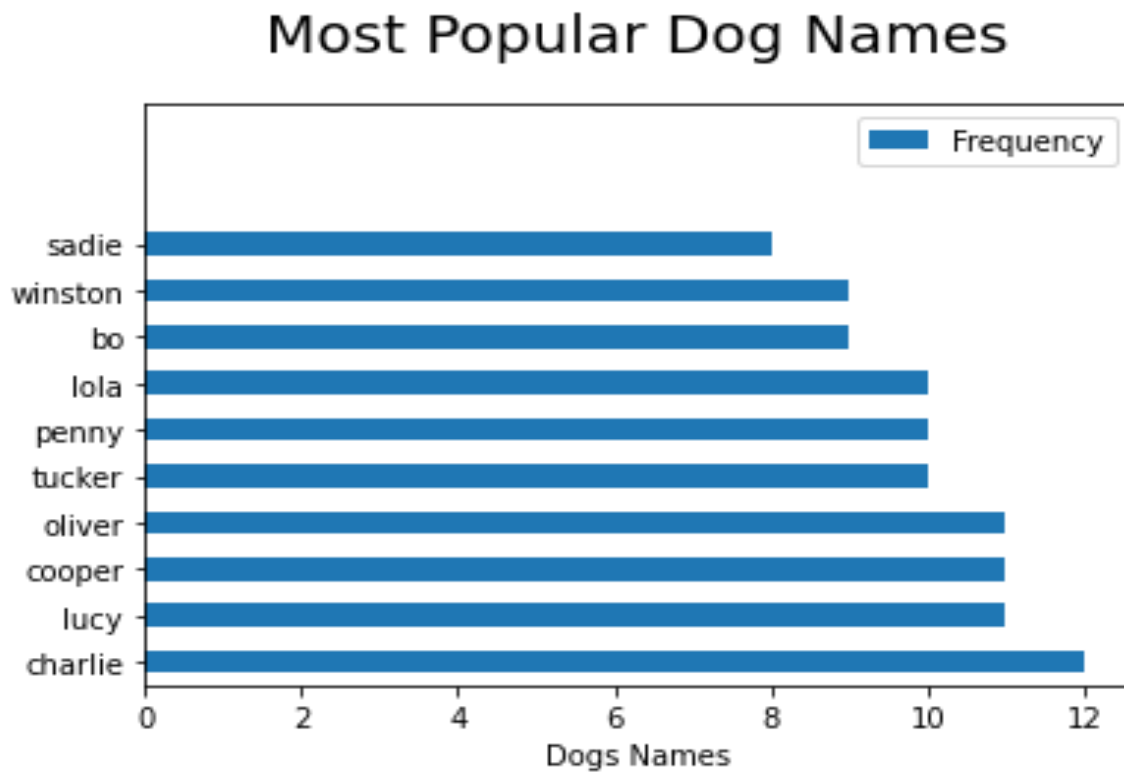
Which dog breed is the most popular, Which dog names are more popular among dog lovers, these questions and more were answered in the analysis below;

*Insights No.1:* Which dog stage is more popular in the dataset?



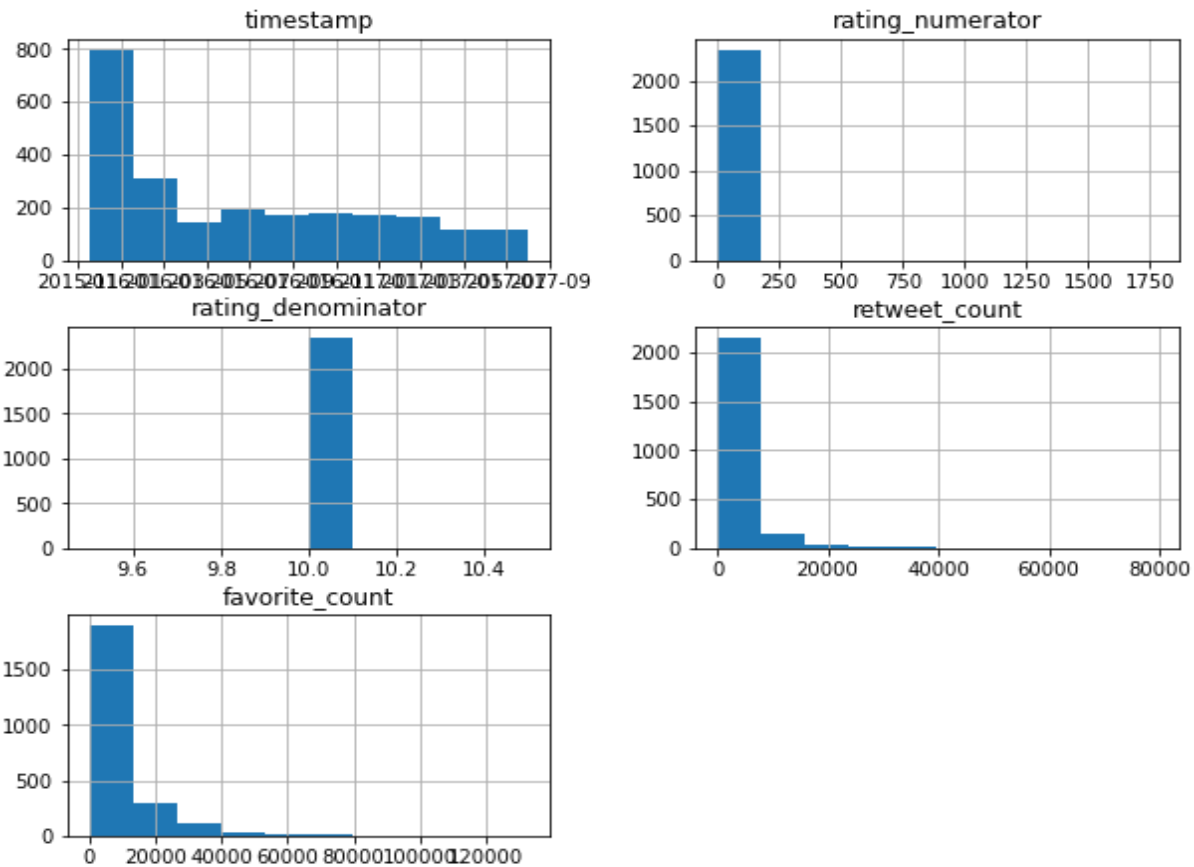
From the above bar chart, we can see that pupper is the most popular dog stage while only a handful of dogs are in the floofer stage.

*Insight No.2:* Which dog names are popular among dog owners?



Interesting insights from this horizontal bar chart shows that the name Charlie is the most popular dog name, closely followed by Lucy, Cooper and Oliver.

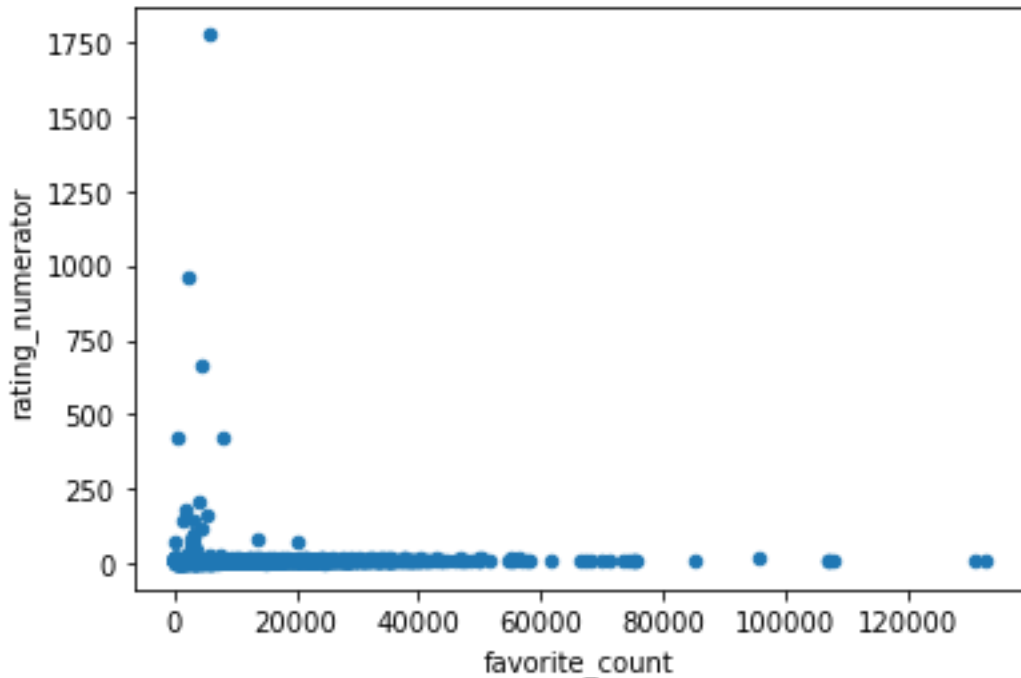
*Insight No.3:*



Time stamp is skewed to the right as we have more tweets from 2015 and 2016, retweet count is also skewed to the right with more retweets ranging from 1- 5000, and just a handful from 5000- 20000 and a few around 40000-70000

This is also the case for favorite count. Rating denominator is different because it consist of the same value(10)

*Insights No.4:* What is the relationship between dog ratings and favorite count?



From the scatterplot above we can see that there is no correlation between the two as they are both moving in different directions.

*Insights No.5:* What is the most occurring dog breed in the dataset?

```
# check value counts for dog breed
twitter_archive_master['breed'].value_counts()[:5]
```

```
: nan                323
golden_retriever     165
labrador_retriever   116
pembroke              94
chihuahua            92
Name: breed, dtype: int64
```

The top dog breeds here as seen above are the golden retriever, labrador retriever, pembroke and the chihuahua. These breeds are popular among dog lovers.

*Insights No. 6:*

Out[61]:

	name	dog_stage	retweet_count	favorite_count	jpg_url	img_num	breed
0	NaN	puppo	48265	132810	https://pbs.twimg.com/media/C2tugXLXgAArJO4.jpg	1	lakeland_terrier

[62]:

```
twitter_archive_master[twitter_archive_master.retweet_count == twitter_archive_master.retweet_count
```

Out[62]:

	name	dog_stage	retweet_count	favorite_count	jpg_url	img_num	breed
	NaN	doggo	79515	131075	https://pbs.twimg.com/ext_tw_video_thumb/74423...	1	labrador_retriever

Unfortunately, this favorite lakeland\_terrier (puppo) does not have a name but has the highest favorite count of 131075

The dog with the most retweets is a labrador\_retriever (doggo) with no name too, He has a retweet count of 79515

This also goes to prove the insight above that a dog got the most retweets does not mean he is the most favorite.

THE END

Thank you.