

We Rate Dogs:

Research of the principal trends in Twitter for @WeRateDogs

Iván Munoz Nunez (1)

1: A pseudorandom guy.

Abstract

With the development of data science, and the rapid broadcasting of his benefits for companies, universities, and NGO's, acknowledge and domain of their fundamental skills has become indispensable to be linked with the requirements they have today. This report condenses an important amount of data related to cute, hilarious, and priceless dogs and how their owners interact with others on Social Media. Relevant stats have been obtained from this data analysis process, involving gathering, cleaning, and assessing data: the most popular breeds, most tweeted, and most replied dogs. The author is sure this study case is the tenderest he had ever seen.

Principal Findings

It was necessary to create a strategy for wrangling data, since all data sources had unordered, unformatted, and uncomplete data. But the love for the dogs is greater.

Dog-olitic Analysis

The most popular breeds are described in the attached figure:

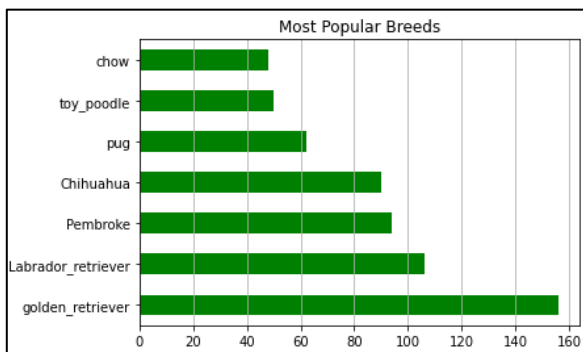


Figure 1 - Most popular dog Breed

This Twitter profile (@WeRateDogs) is well-known for classifying their protagonists by stages. Those are described in Figure 2:

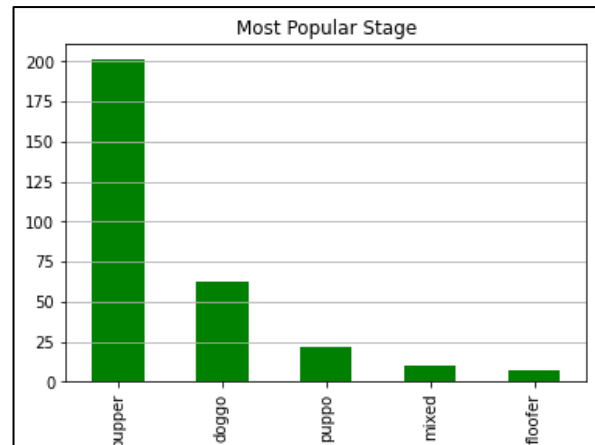


Figure 2 - Most popular dog Stages

As reader can see, **"pupper"** is, by far, the most popular stage between @WeRateDogs Twitter.

The most retweeted dogs (top 3) are described in Appendix I.

The reader can find most *"favorited"* dogs (top 3) are in the Appendix II.

The analysis for the most popular names is described as follows (assuming dogs are uploaded once only):

Udacity Data Analysis Fundamentals - Project 4

Data Wrangling

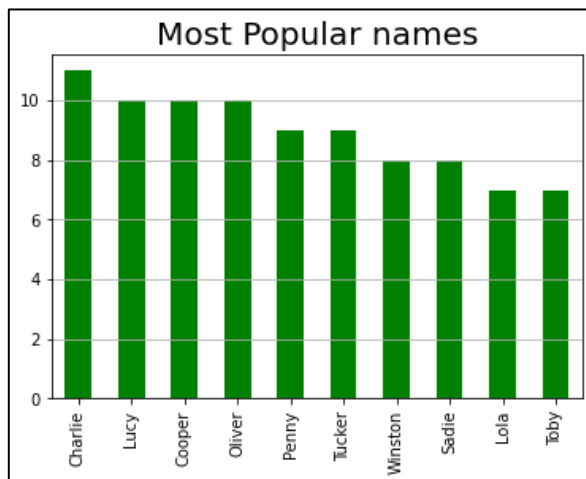


Figure 3 - Most popular dog Names

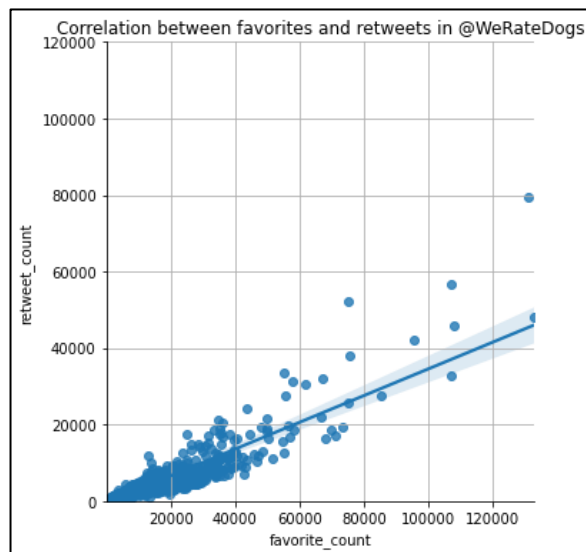


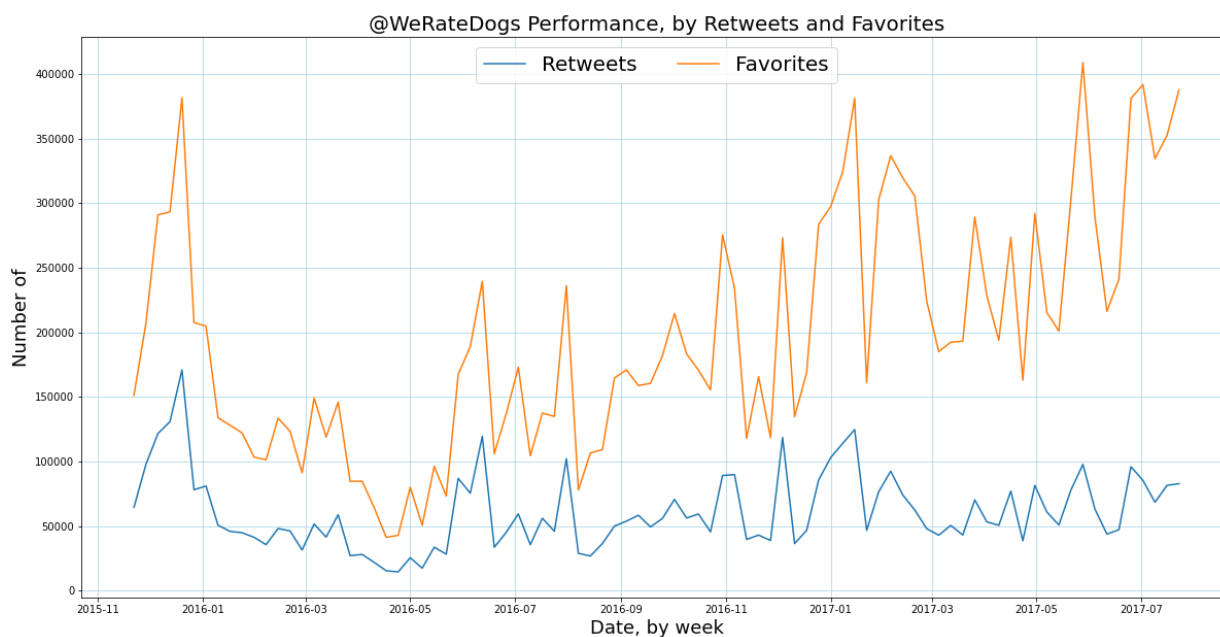
Figure 4 -Retweets vs Favorites scatter plot.

Retweets and Favorites: What do they say to us?

Does exist a relation between favorites and retweets? The next figure indicates that, in fact, exists:

This face-to-face analysis has a strong correlation ($r^2 = 0.913$) for linear regression.

The page was analyzed in its performance as well. This, to study the trend between November 2015 and July 2017. The following figure describes performance based on retweets and “likes” for this period:



Appendix (this is the best part):

Appendix I: Most retweeted dogs



Retweets: 56.625

Name: Stephan

Tweet: This is Stephan. He just wants to help. 13/10 such a good boy



Retweets: 79.515

Name: Unknown

Tweet: Here's a doggo realizing you can stand in a pool. 13/10 enlightened af (vid by Tina Conrad)



Retweets: 52.360

Name: Unknown

Tweet: Here's a doggo blowing bubbles. It's downright legendary. 13/10 would watch on repeat forever (vid by Kent Duryee)

Udacity Data Analysis Fundamentals - Project 4
Data Wrangling

Appendix II: Most Liked (Favorite) dogs



Favorites: 131.075

Name: Unknown

Tweet: Here's a doggo realizing you can stand in a pool. 13/10 enlightened af (vid by Tina Conrad)



Favorites: 132.810

Name: Unknown

Tweet: Here's a super supportive puppo participating in the Toronto #WomensMarch today. 13/10



Favorites: 107.956

Name: Duddles

Tweet: This is Duddles. He did an attempt. 13/10 someone help him (vid by Georgia Felici)