

Act_report:

Insight 1:

In this analysis, I found golden_retriver holds the first position with 1431940 favorite counts and Labrador_retriever is in the second position with 880177, followed by Pembroke, with 846279 and chihuahua with 584426. 10th and 9th position hold by Pomeranian and cocker_spaniel with 258530, 261801. The top 4 favorite dogs have the top rank too but the rank is given with some exceptions like a ranking_numerator is way higher than the denominator in some places. Some rows have values like 1776.0, 666.0, and 420.0 which is way higher than the top rank here. It can skew the data.

Insight 2:

Here you can find that the puppies are the most rated dogs after that comes doggo, and puppo. floofer are the least rated. But in most cases dog's stage is not given at all, the column holds the null values. pupper count in data is 224 and the rating sum is 2405.27 after pupper comes the doggo whose count is 87 and the rank sum is 1024. there is only 9 floofer with 107 total sums of ranks.

Insight 3:

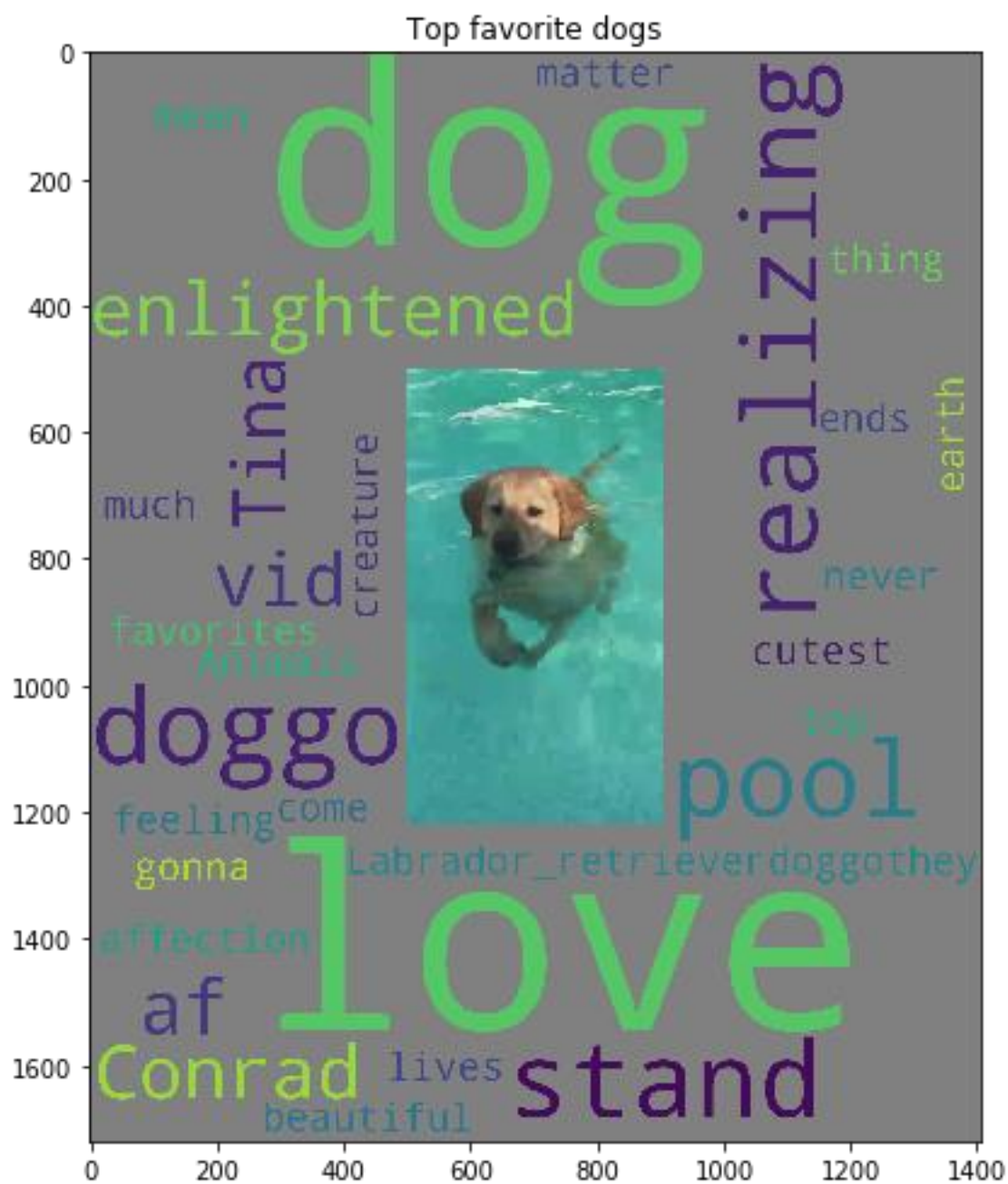
In retweet count, the top position is held by the golden retriever with 400242 retweets same as in the favorite count table. In fact, the top 5 position is the same as in the favorite table held by Labrador_retriever, Pembroke, chihuahua, and Samoyed. But in 10th position, Eskimo_dog came as a new entry with rank 212 and 78171 retweets.

Visualization:

Here I use text and image data to visualize the top five favorite dogs with wordcloud package. I use their images and create wordcloud around them using columns 'text', 'pre_breed', 'stage', and text that explain a dog pretty well.

Here is the visualization for Top 5 favorite dogs:

- 1.



2.



3.



4.



5.

Top favorite dogs

