

## Relatório da Análise e Visualização

Ao iniciar a análise, primeiramente fiquei intrigado sobre quais as raças que possuíam maior número de posts, ou seja, a que o WeRateDogs mais opta por postar conteúdo e obtive a seguinte relação, buscando pelas três raças com maior quantidade de posts:

Previsão de raça	Nota Média	Quantidade de posts	Número médio de Retweets	Número médio de Favoritadas
Golden Retriever	1.156774	155	3677.603774	12166.387097
Labrador Retriever	1.119811	106	3599.645161	11795.622642
Pembroke (Corgi)	1.142553	94	2951.585106	10801.404255

E o resultado foi proporcional ao que encontramos na internet hoje de páginas de redes sociais que postam conteúdo desse tipo, vários e vários posts com Goldens, Labradores e Corgis.

Fiquei curioso para saber se as três raças também estariam na relação das que recebem mais curtidas e retweets, até porque para haver uma quantidade tão grande de postagens destas raças, a repercussão deve ter sido positiva.

Primeiramente extrai os dados das três raças que receberam mais retweets em média, e obtive o resultado da tabela a seguir:

Previsão de raça	Número médio de Retweets	Quantidade de posts
Bedlington Terrier	7872.333333	6
Afghan Hound	5633.000000	3
Standard Poodle	5213.272727	11
French Bulldog	5160.838383	30
English Springer	5114.500000	10

E para a minha surpresa, nenhuma das raças com maior quantidade de posts apareceu, e podemos notar que as duas que tiveram a média mais alta, tiveram uma quantidade relativamente baixa de posts, porém a média está bem acima da média dos três com mais postagens.

E por curiosidade quais os três posts que tiveram maior quantidade de retweets e as raças previstas neles?

Tweet	Previsão de raça	Número de Retweets	Confiabilidade da Previsão
<a href="#">744234799360020481</a>	Labrador Retriever	83587	82,53%
<a href="#">739238157791694849</a>	Eskimo Dog	61909	50.33%
<a href="#">807106840509214720</a>	Chihuahua	60938	50.53%

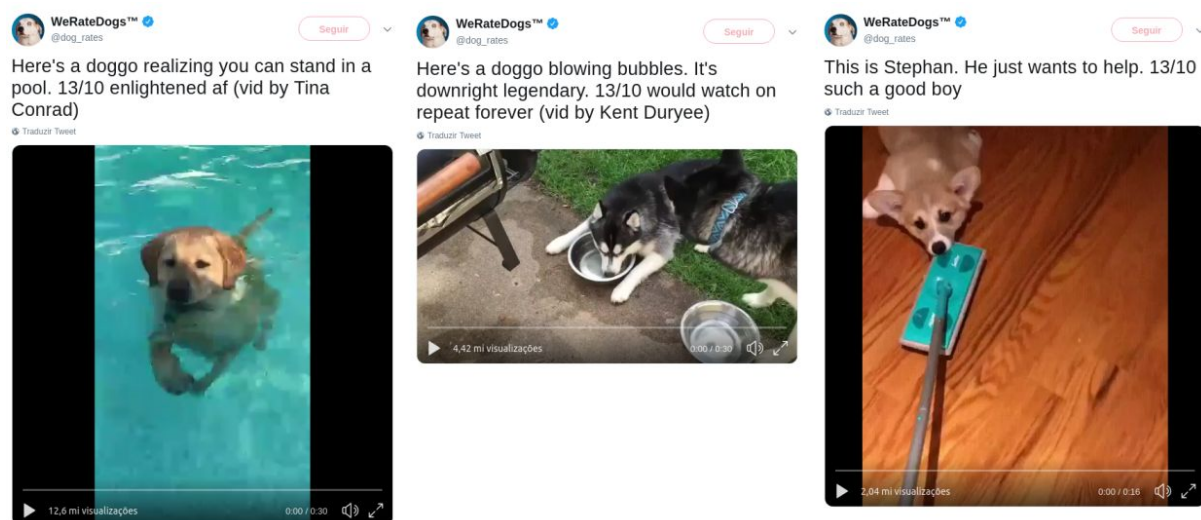


Imagem 1: Os três tweets com maior número de Retweets.

Tivemos então um Labrador Retriever que tiveram incríveis 83587 retweets, seguido pelo Eskimo Dog (Que apesar de ser um Husky, tem muita semelhança com um Eskimo Dog) com 61909 e pelo Chihuahua (Que na verdade é um Pembroke e a previsão está errada) com 60938, coincidentemente, os três bons garotos com nota 13/10. Vale notar também que os três tweets tem vídeos anexados, e não imagens, este poderia ser um bom indicativo para maior alcance de retweets.

Ao extrair a maior quantidade de favoritas por tweets estava esperando um resultado parecido com o do número de retweets, e então obtive o resultado mostrado na seguinte tabela:

Previsão de raça	Número médio de Favoritadas	Quantidade de posts
Bedlington Terrier	24135.833333	6
Saluki	23205.000000	4

French Bulldog	19619.733333	30
Bouvier Des Flandres	17264.000000	1
Afghan Hound	16540.666667	3

E como esperado, 3 das 5 raças que estavam no top 5 do maior número de retweets também apareceram no maior número de favoritadas. As novas raças que surgiram, também possuem uma baixa contagem de tweets, o que mostra que tiveram tweets isolados destas raças com um grande número de acesso. Talvez seja uma boa ideia para a página fazer mais posts destas raças que atingiram um número tão grande de retweets e favoritadas.

E vamos então também extrair os três posts com maior número de favoritadas:

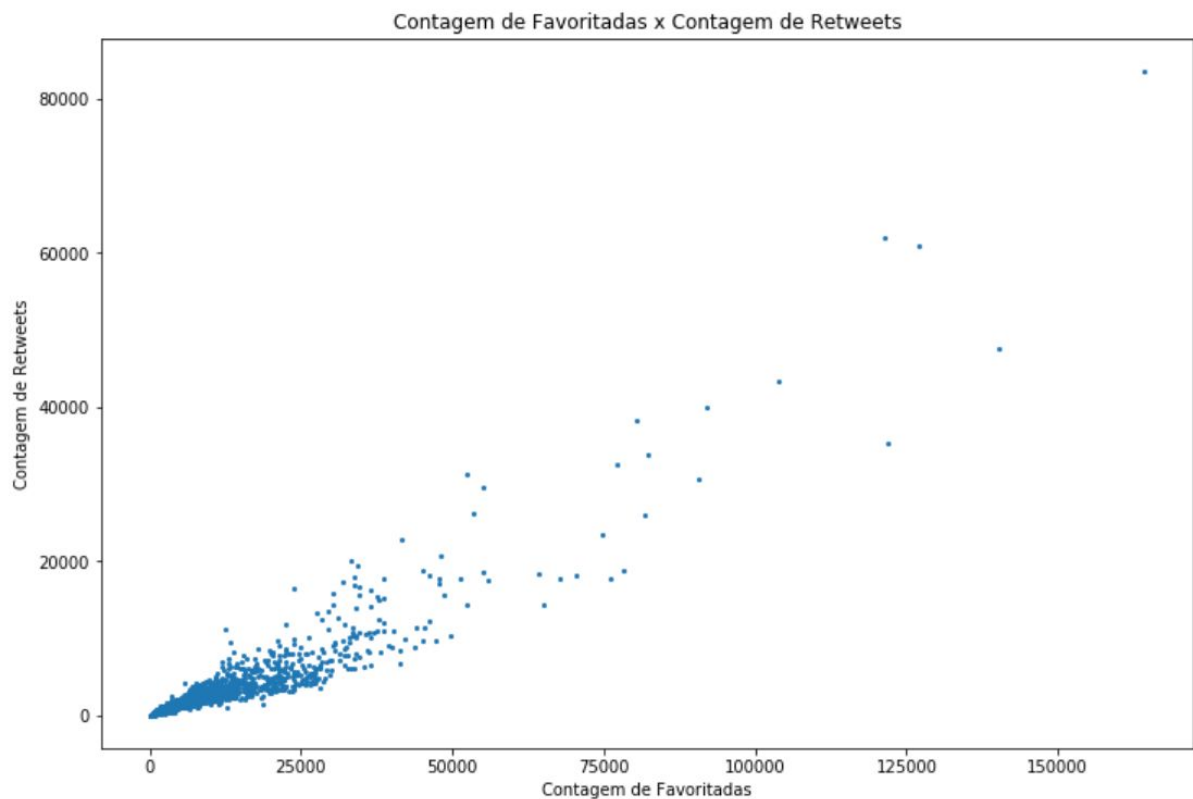
Tweet	Previsão de raça	Número de Favoritadas	Confiabilidade da Previsão
<a href="#">744234799360020481</a>	Labrador Retriever	164188	82,53%
<a href="#">822872901745569793</a>	Lakeland Terrier	140314	19.60%
<a href="#">807106840509214720</a>	Chihuahua	126948	50.53%



Imagem 2: Os três tweets com maior número de favoritadas.

E aqui encontramos novamente o Labrador Retriever e o Chihuahua (que na verdade é um corgi), mantendo suas respectivas colocações, e em segundo, temos um novo tweet de um Lakeland Terrier (que também não é um Lakeland Terrier), e que desta vez não foi um video, mas uma imagem referente a uma hashtag que estava em alta [#WomensMarch](#). Fazer postagens sobre temas em alta com imagens de cachorros fofos aparentemente também gera grande repercussão.

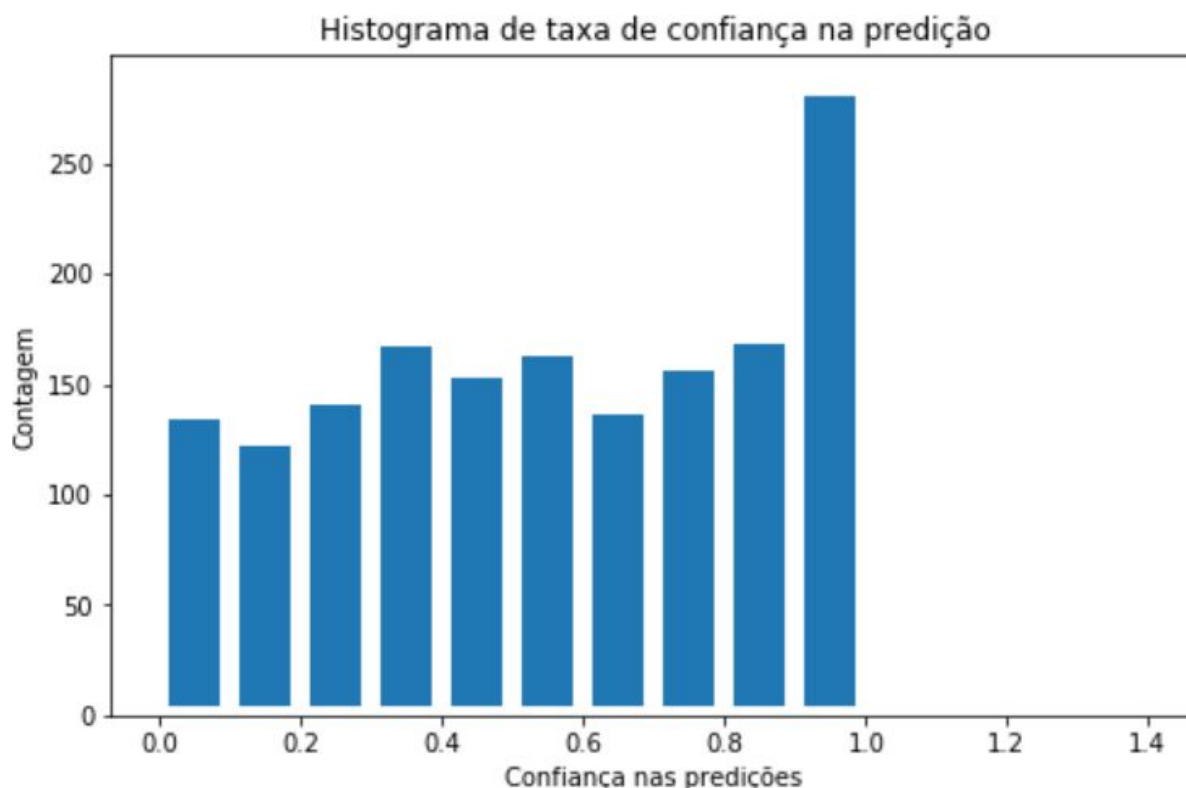
Em meio a estas análises foi perceptível que o número de retweets e favoritadas estão de certa forma diretamente relacionados, coloquei então uma visualização dos dois parâmetros para observar o comportamento.



*Imagem 3: Gráfico da relação de contagem de favoritadas por retweets, cada ponto azul no gráfico é a representação de um tweet, e suas coordenadas suas respectivas contagens.*

E assim com todos os dados colhidos apresentados podemos verificar de fato que quanto maior a quantidade de favoritadas, a tendência é que a contagem de retweets também seja alta e vice-versa. Isso se dá ao fato de que quando pessoas dão retweet, a postagem aparece no feed de outras que a seguem e trazem maior atenção para favoritar ou também retweetar, da mesma forma quando algum contato próximo seu, ou um grupo de pessoas que você segue favoritam algo, uma notificação também é gerada para atrair sua atenção para aquele tweet.

Em relação à confiabilidade observamos também que a taxa de confiabilidade estava baixa para os casos analisados isoladamente, tracei então um histograma das confiabilidades encontradas.



*Imagem 4: Histograma da taxa de confiança, quantidade de tweets que tiveram a taxa de confiança relativa a proporção apresentada nos números do eixo x.*

Como resultado obtivemos uma variação muito grande de confiabilidade, além dos 304 tweets que não apresentaram reconhecimento algum de imagem. Então a previsão apresentada nos tweets não tem uma confiabilidade muito alta, os tweets com uma confiabilidade relativamente alta, de acima de 80% de confiança representam apenas 458 de um total de 1967 tweets, ou seja 23,29% da base de dados.

Por fim gostaria apenas de relatar que utilizei as raças para fazer a análise pois 84,54% da base de dados tinham alguma previsão de raça (mesmo que baixa), uma análise para o tipo de cachorro apresentado na imagem (doggo, floofer, puppo e pupper) tornou-se inviável pois poucas linha tinham dados pertinentes sobre a existência destas classes nas imagens.