

AVALIANDO CACHORROS!

Oi pessoal!

Que tal analisarmos as avaliações de cachorros feitas pelo usuário @dog_rates no tweeter?! Ele posta fotos de diferentes cachorros e dá nota para eles!

Tem muitas fotos interessantes, vejam essas aqui:

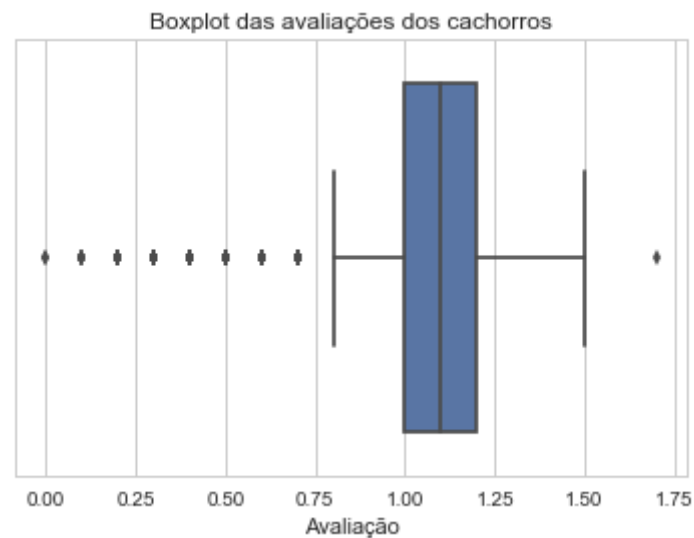


Bem legal, né?!

As avaliações que ele dá são mais ou menos essas: 12/10, 13/10, 14/10, etc.

Mas é complicado comparar os números dessa forma, então iremos realizar a divisão. Fazendo do nosso jeito as avaliações acima ficariam assim: 1.2, 1.3, 1.4, etc.

Pois bem, vejam só como fica a distribuição das frequências de todas as avaliações que esse usuário fez até a data de 01/08/2017:



Esse gráfico mostra que as avaliações variam de 0 a 1.7. E pelo menos metade dessas avaliações deve estar entre 1.0 e 1.2.

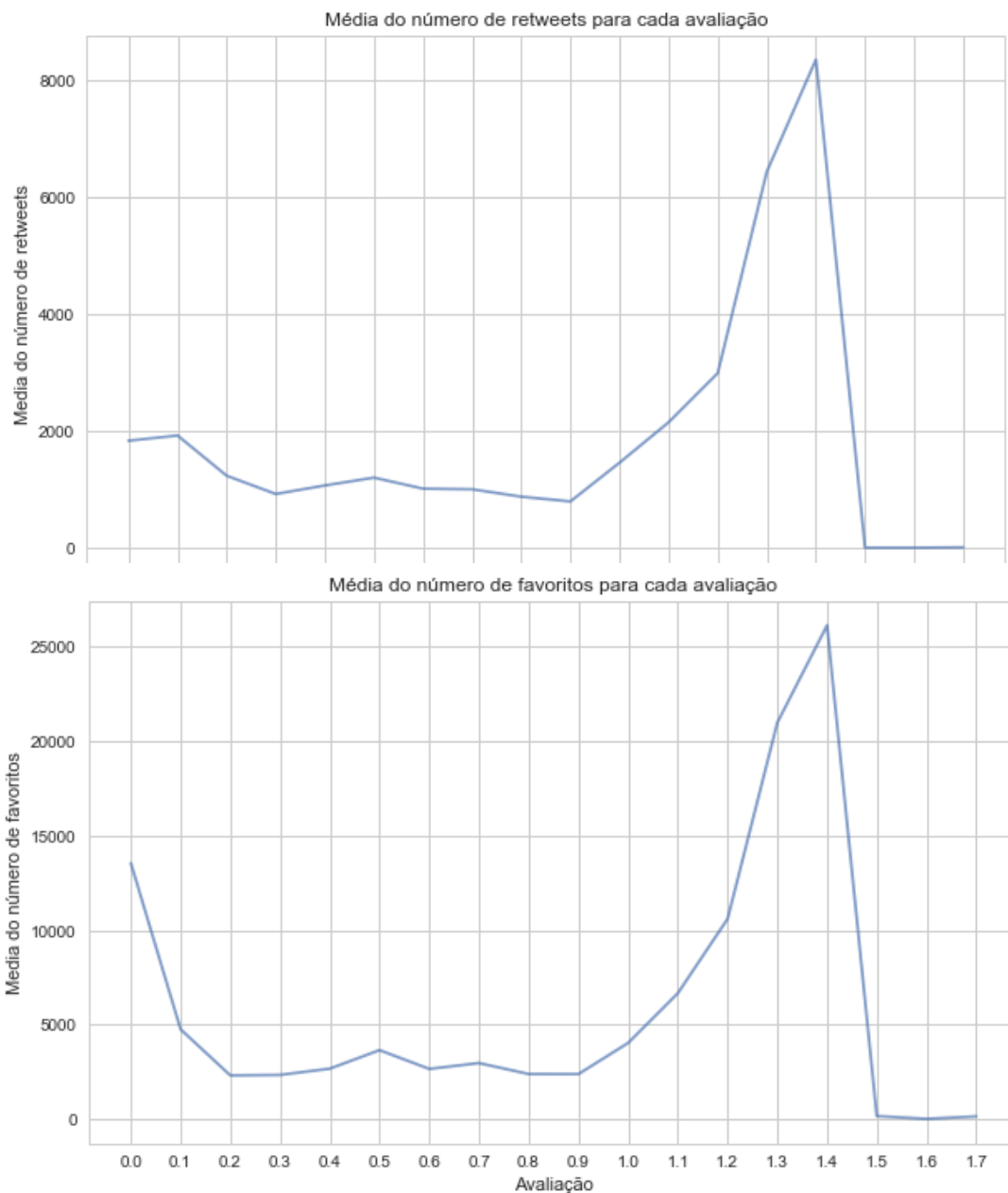
Poderíamos nos perguntar: será que as postagens com muitos retweets possuem também, em geral, muitos favoritos?? SIM!!! É o que observamos no gráfico a seguir:



Esse gráfico mostra também que as postagens com números muito elevados de retweets e de favoritos são poucas, se compararmos com a quantidade daquelas que foram menos retweetadas e favoritadas.

A AVALIAÇÃO DOS CACHORROS E O NÚMERO DE RETWEETS E FAVORITOS DA RESPECTIVA POSTAGEM

Outra questão importante: a avaliação dos cachorros está relacionada, de alguma forma, ao número de retweets e de favoritos? Podemos encontrar a resposta observando os seguintes gráficos:



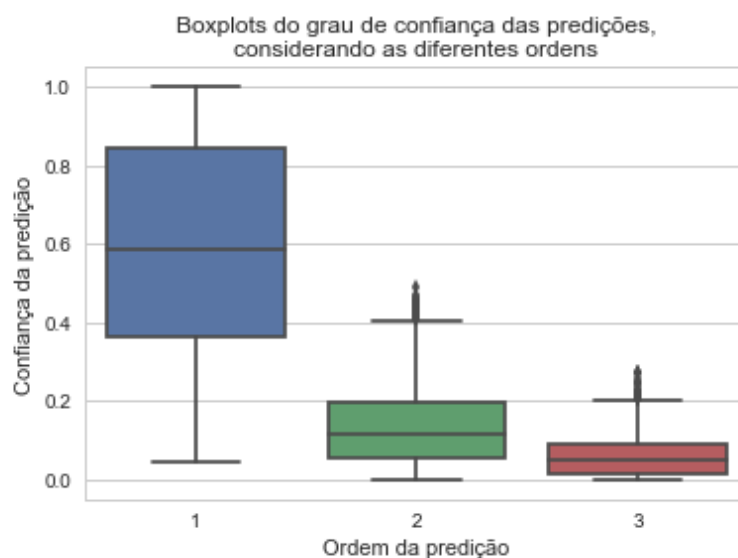
Que gráficos interessantes! Vejam como eles se parecem um com o outro! Eles sugerem que, quando a avaliação do cachorro é extremamente baixa (0.0 ou 0.1), os números de retweets e de favoritos são, em média, relativamente elevados. Mas, aumentando apenas um pouco a avaliação, o número médio de retweets e de favoritos já cai drasticamente.

Depois das avaliações de nota zero, os gráficos das avaliações seguintes mantêm mais ou menos uma constância de poucos retweets e favoritos, voltando a crescer esse número quando as avaliações superam a nota de '0,9'. O gráfico é crescente a partir daí, atingindo seu pico na avaliação de '1,4', e depois decresce drasticamente. Esse comportamento sugere que existe uma relação entre a avaliação do cachorro e o número de retweets e de favoritos. Essa relação é mais evidente no intervalo de avaliações de 0,9 a 1,4, no qual os gráficos das médias de retweets e de favoritos são estritamente crescentes.

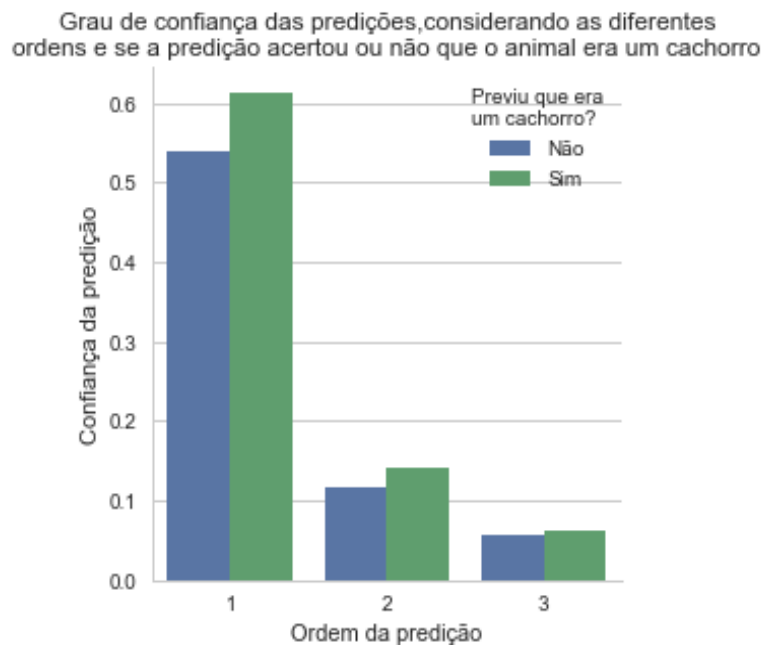
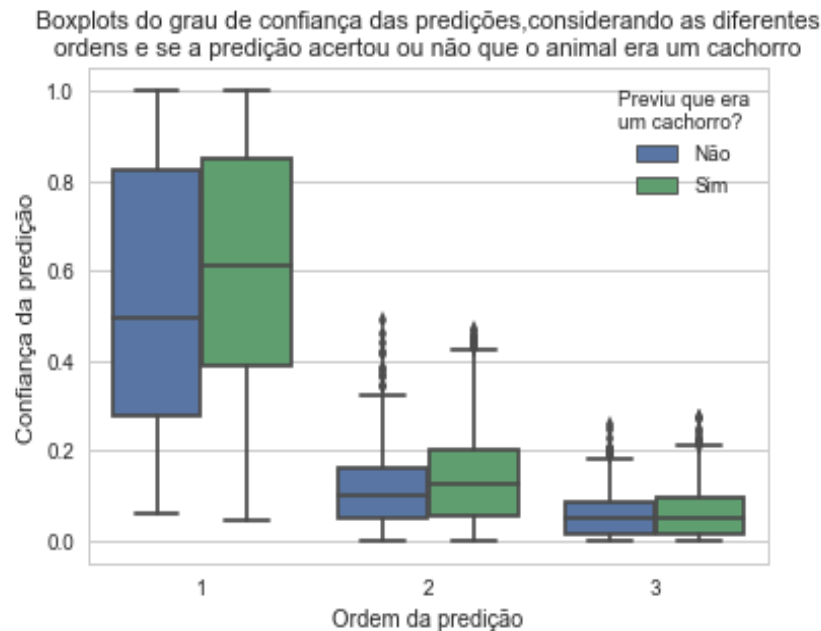
O ALGORÍTIMO DE PREDIÇÃO

Existe um algoritmo que prevê, com base em uma imagem, qual é esse animal e a sua raça. Ele traz, para cada imagem, três predições com diferentes graus de confiabilidade, sendo a predição 1 a mais confiável e a predição 3 a menos confiável.

Utilizando esse algoritmo nas imagens publicadas pelo usuário @dog_rates, como será que fica o grau de confiança das predições? É o que vamos analisar a seguir.

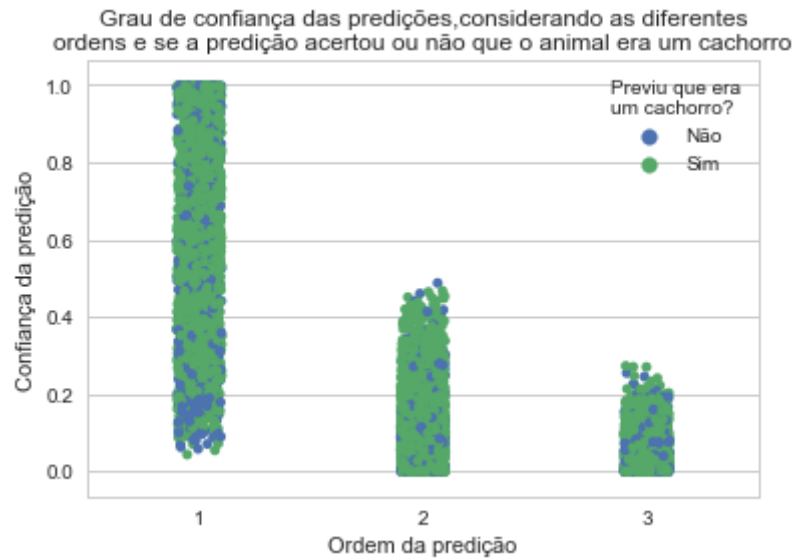


Vejam nesse gráfico como a predição 1 é realmente bem superior às demais!! Agora, por melhor que sejam as três predições do algoritmo, todas as três erraram, em vários casos, até o tipo de animal. Não conseguiram concluir que a imagem se tratava de um cachorro. Vajamos, nos gráficos, abaixo como fica a distribuição das que acertaram e das que erraram o tipo de animal:



Esses gráficos mostram que, em geral, para todas as três ordens de predição, o grau de confiança das predições que acertam o tipo de animal costuma ser um pouco superior ao grau de confiança daquelas que erram.

Apesar de tudo, em relação ao total de previsões de cada ordem, são poucas aquelas que erram o tipo de animal. Podemos ver isso no gráfico a seguir:



A quase totalidade dos gráficos está pintada de verde, que apresenta as previsões que acertaram o tipo de animal. Apenas uma pequena parte está pintada de azul, que representa as previsões que erraram o tipo de animal.

Gostaríamos de saber também se a avaliação dos cachorros está diretamente relacionada ao grau de confiança das três previsões do algoritmo. Para responder esse questionamento, vamos ao gráfico:



Percebam como essas três linhas são irregulares! Não temos como afirmar que a previsão será mais confiável quando estivermos diante de notas de avaliação mais altas. Essa relação simplesmente não existe!!!

OS ESTAGIOS DE VIDA DOS CÃES

Em algumas postagens o usuário @dog_rates, além de colocar a avaliação do cachorro, também colocou o seu estágio de vida. São quatro os estágios que ele coloca, de acordo com o caso: “doggo”, “floofer”, “pupper” e “puppo”. Vejamos algumas fotos para conseguir identificá-los:

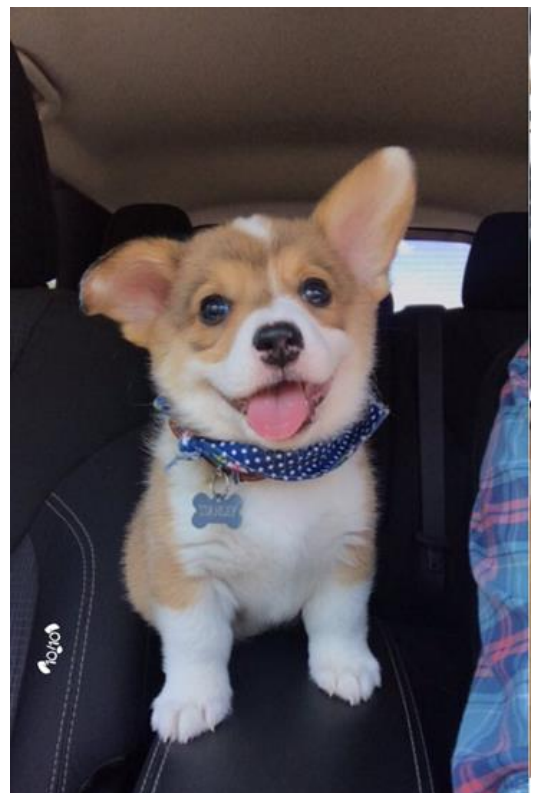
- Doggo:



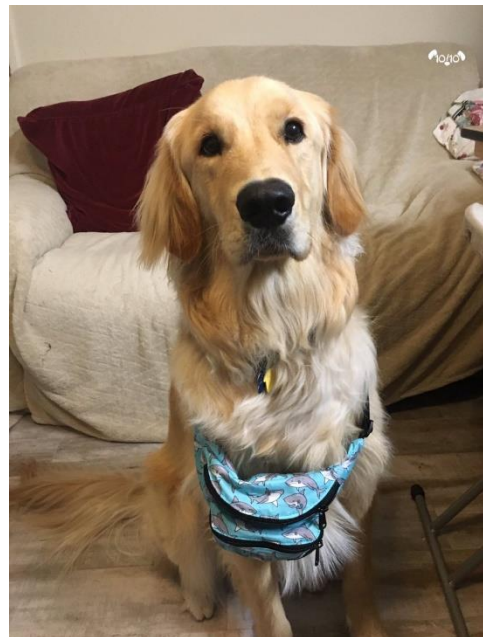
- Floofer:



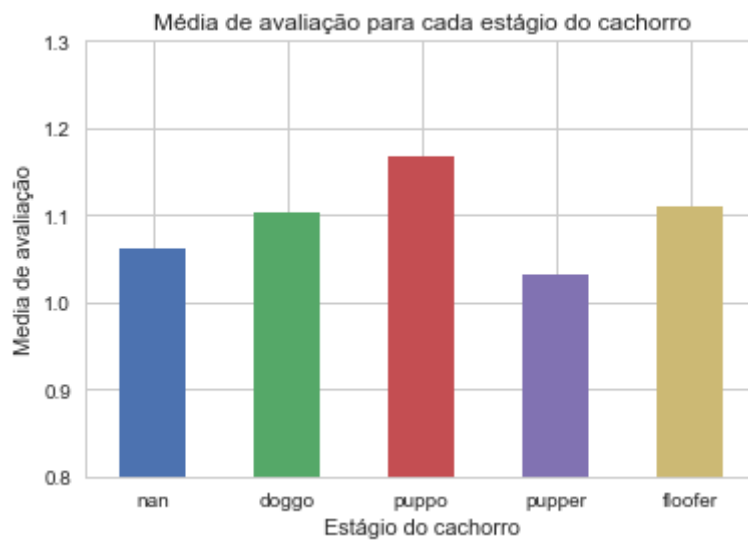
- Pupper:



- Puppo:

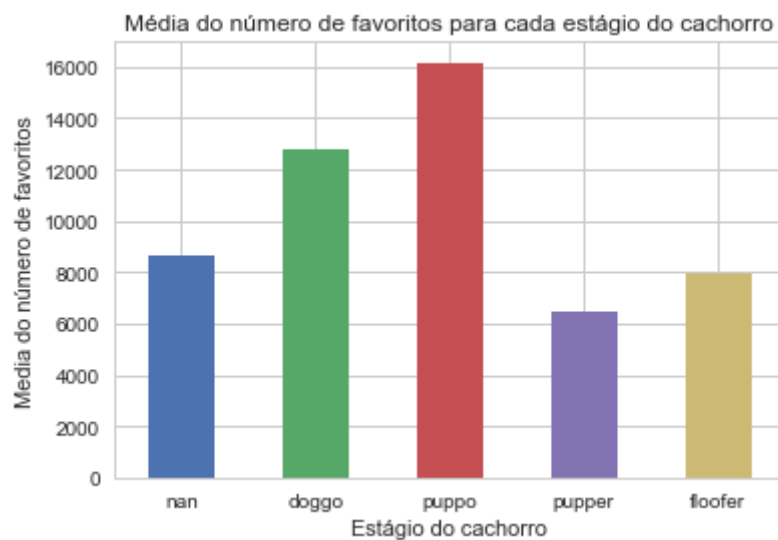
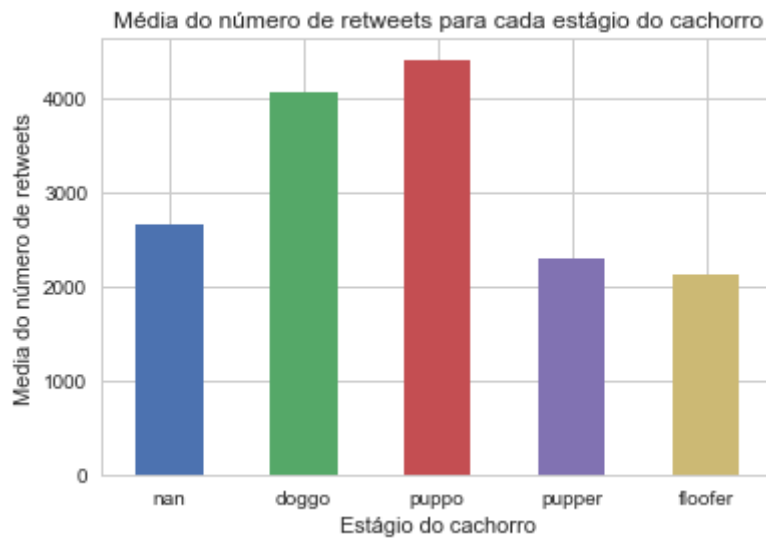


Que bacana! E para qual desses estágios será que o usuário @dog_rates costuma dar as melhores avaliações??? Vamos ver:



Esse gráfico mostra que os cães no estágio “puppo” recebem, em média, avaliações mais altas que os demais. Já os cães no estágio “pupper” recebem as piores avaliações, em média.

Mas será que os outros usuários também preferem o estágio “doggo”? Vejamos:



Parece que a resposta é SIM!!! Tanto a média do número de retweets como a do número de favoritos é maior para os cães no estágio “puppo”! Justamente aqueles que recebem, em média, as melhores avaliações!!!

Por hoje é só, pessoal!

Nos vemos na próxima análise!