

Insights

Nome: Leonardo Soares Duarte

Este relatório faz parte do Projeto 2 do Nanodegree Fundamentos de Data Science II – Udacity

1) Porcentagem por Estágio

O gráfico de pizza a seguir mostra os estágios e suas respectivas porcentagens. Podemos ver que, dos 250 tweets com classificação do usuário WeRateDog, 67.2% são Puppies. Segundo o 'Dogtitionary' do próprio autor, puppies são geralmente filhotes. Isso nos mostra uma preferência do usuário em classificar filhotes.

Porcentagem por Estágio

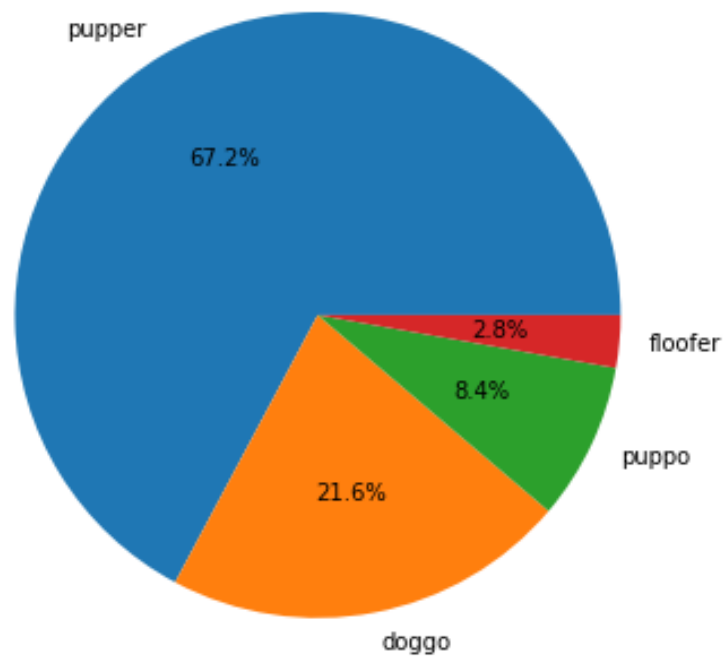


Figura 1

2) 'Favorites' e 'Retweets' por Estágio

Curiosamente vemos que, apesar de ser o estágio com mais classificações pelo usuário, os Puppies foram os que tiveram as menores medianas de 'Favorites' e 'Retweets'. Não encontrei dados que justificassem esse fato.

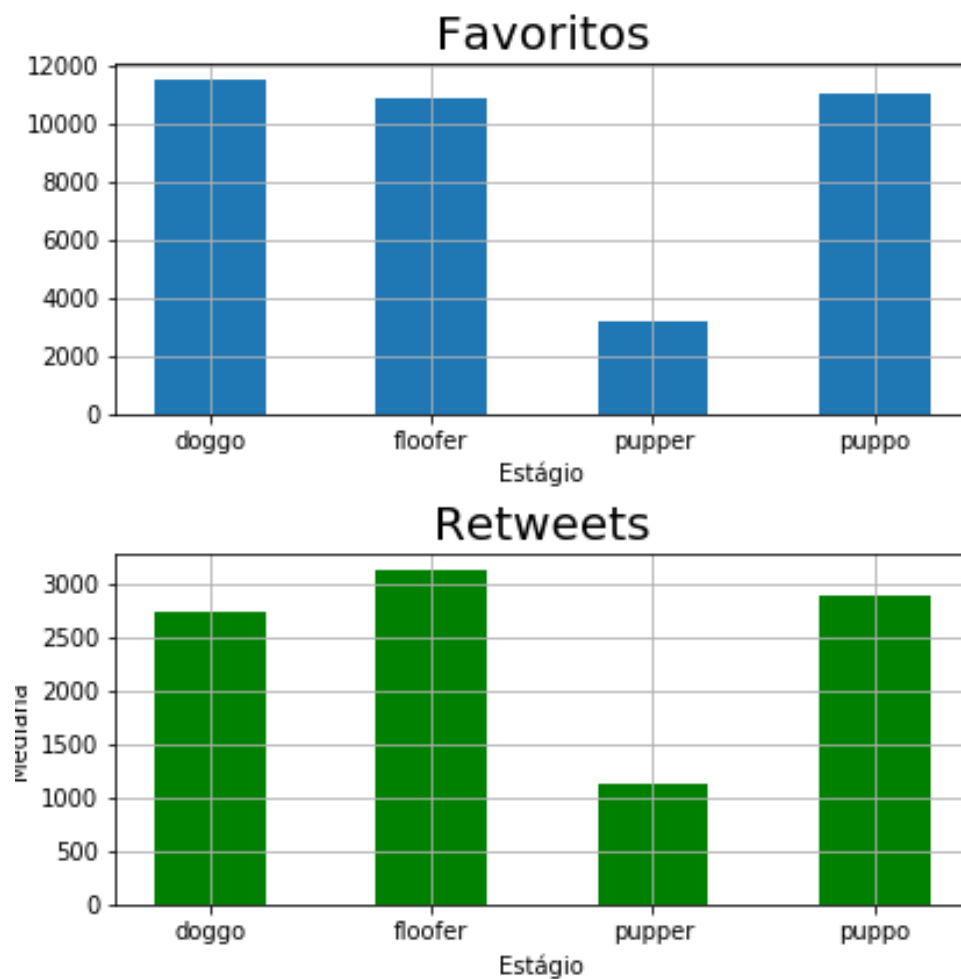


Figura 2

3) Ranking de raças

Já nos gráficos da figura 3 vemos as raças mais populares com base nas medianas de retweets e favoritos. As raças 'Irish Water Spaniel', 'Saluki', 'Flat-Coated Retriever', 'Bouvier des Flandres', 'Black-and-tan Coonhound' e 'Afghan Hound' foram as raças que mais tiveram repercussão entre os seguidores do usuário WeRateDogs. Essas raças estão presentes tanto nos 10 mais favoritos quanto nos 10 mais retweets.

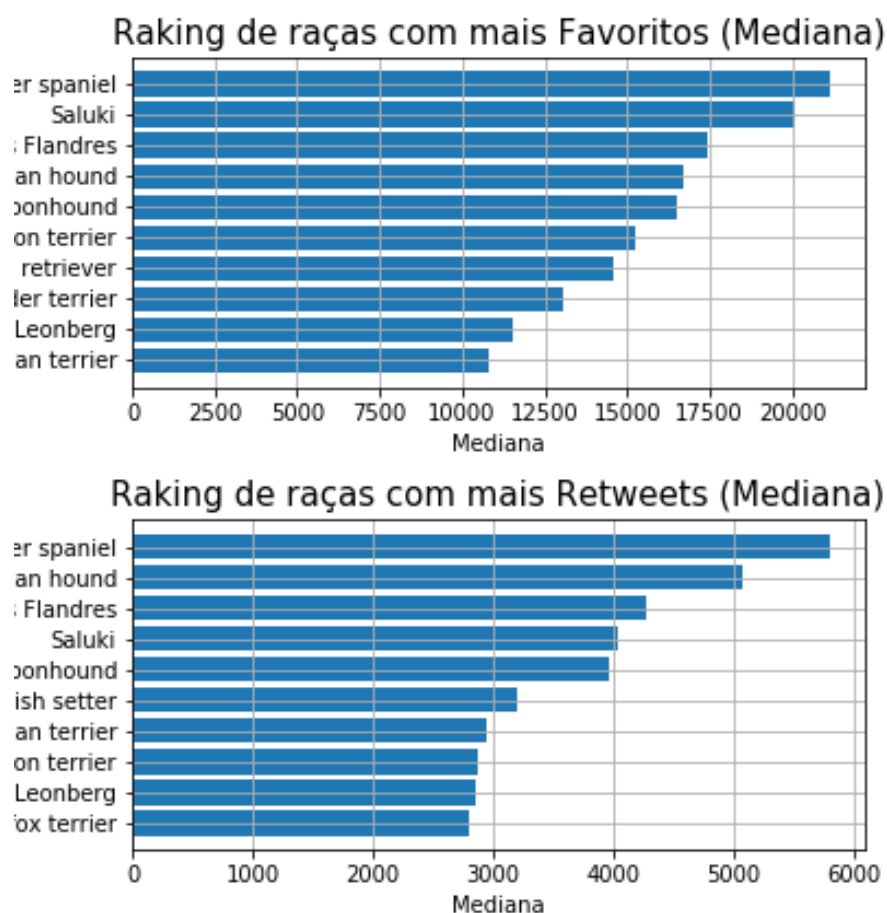


Figura 3

4) Gráfico de dispersão

A figura 4 é um gráfico de dispersão que mostra a quantidade de retweets e favoritos para cada tweet. Nele podemos ver que essas duas variáveis possuem uma relação diretamente proporcional, ou seja, os tweets com mais favoritos geralmente são os com mais retweets.

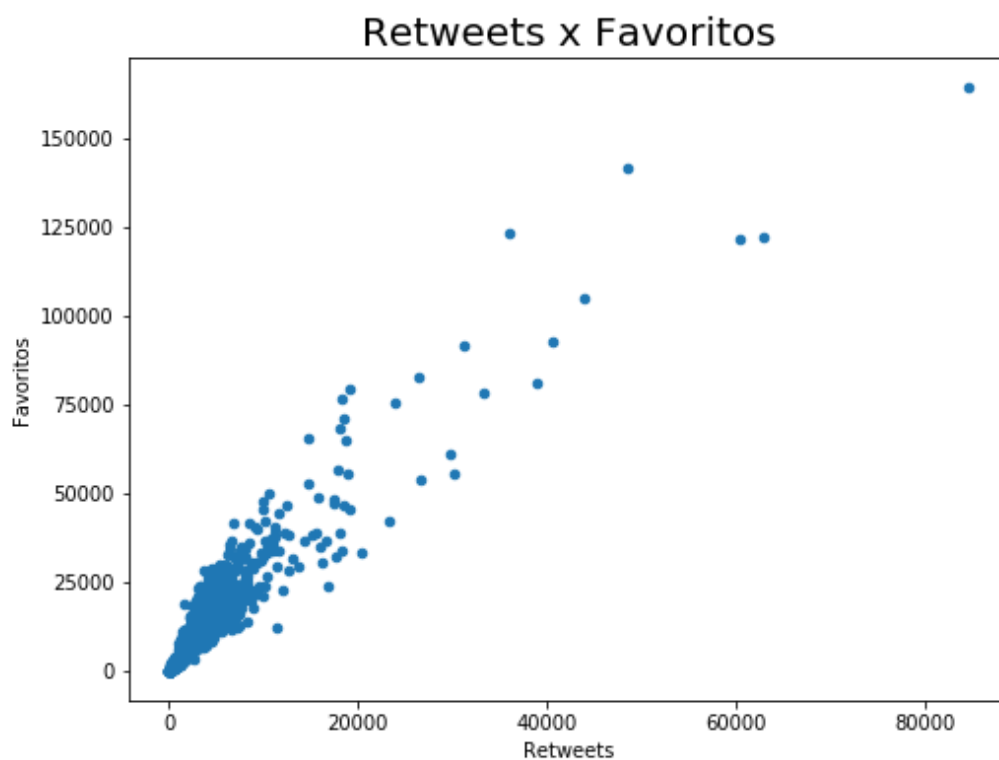


Figura 4