

Apresentação

Este projeto consiste na avaliação dos dados do Arquivo de Tweets do perfil @WeRateDogs, especializado em avaliação e categorização de fotos de cães. Foram mescladas a este arquivo de Tweets informações complementares, como contagem de favoritos e retweets, extraídas da API do Twitter. Como informações adicionais, foram mescladas a predição da raça dos cães presentes nas fotos dos tweets.

O resultado, após a limpeza e união destes conjuntos de dados, é um dataframe com a identificação dos Tweets, sua avaliação, categoria e raça do cão em questão. Assim, posteriormente foram levantadas perguntas primárias sobre os dados, tais como:

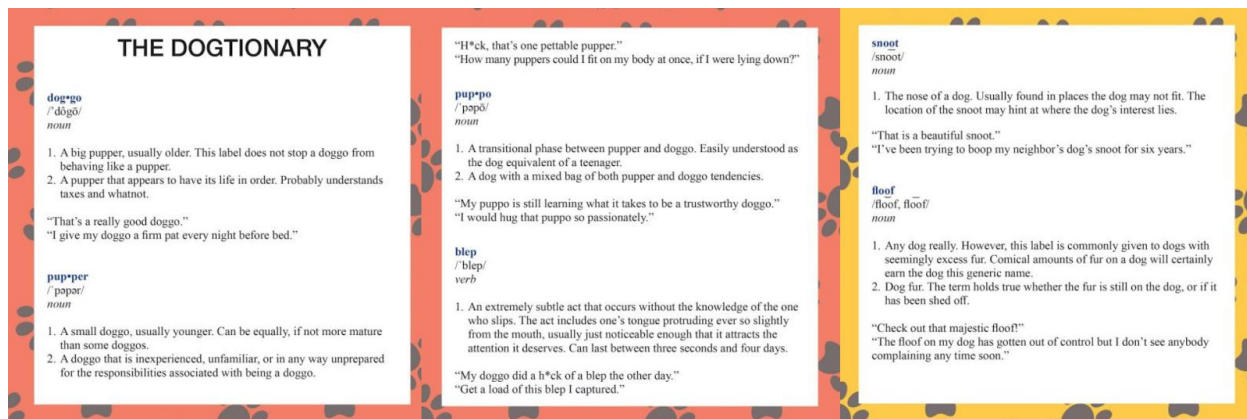
1. Quantos animais foram categorizados?

Dos **1746** registros, **257** foram categorizados, ficando **1489** registros não categorizados.

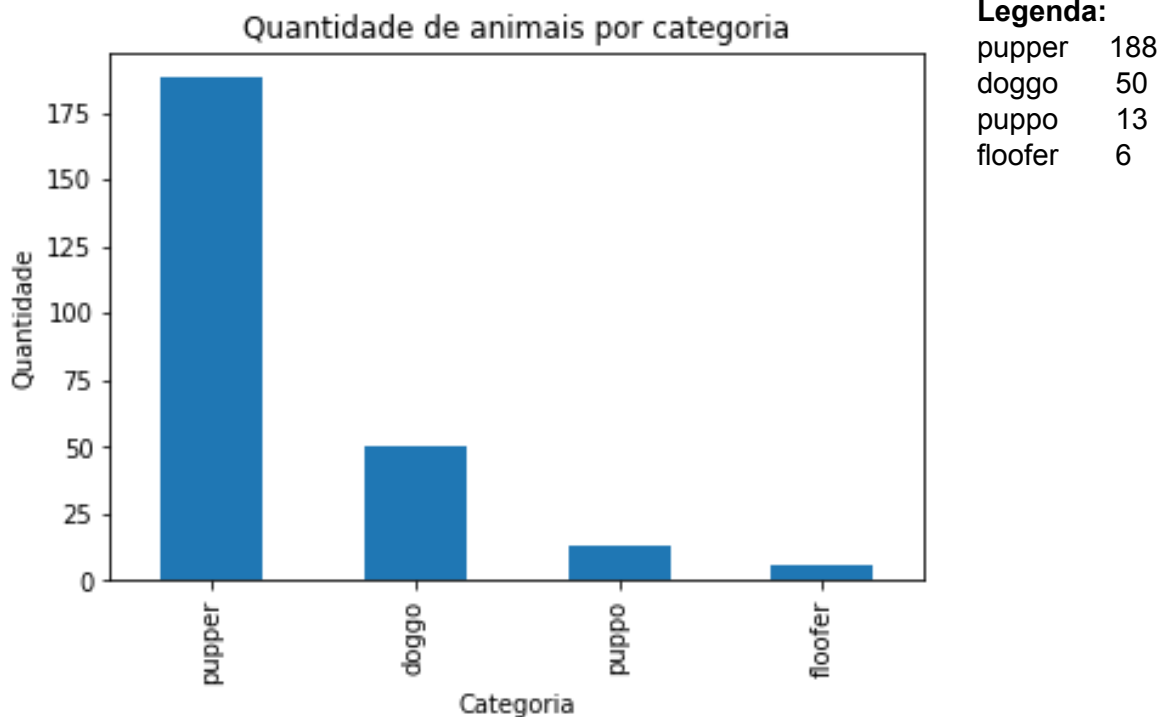
A partir da segunda pergunta foram considerados somente os animais categorizados.

2. Quantos animais há em cada categoria?

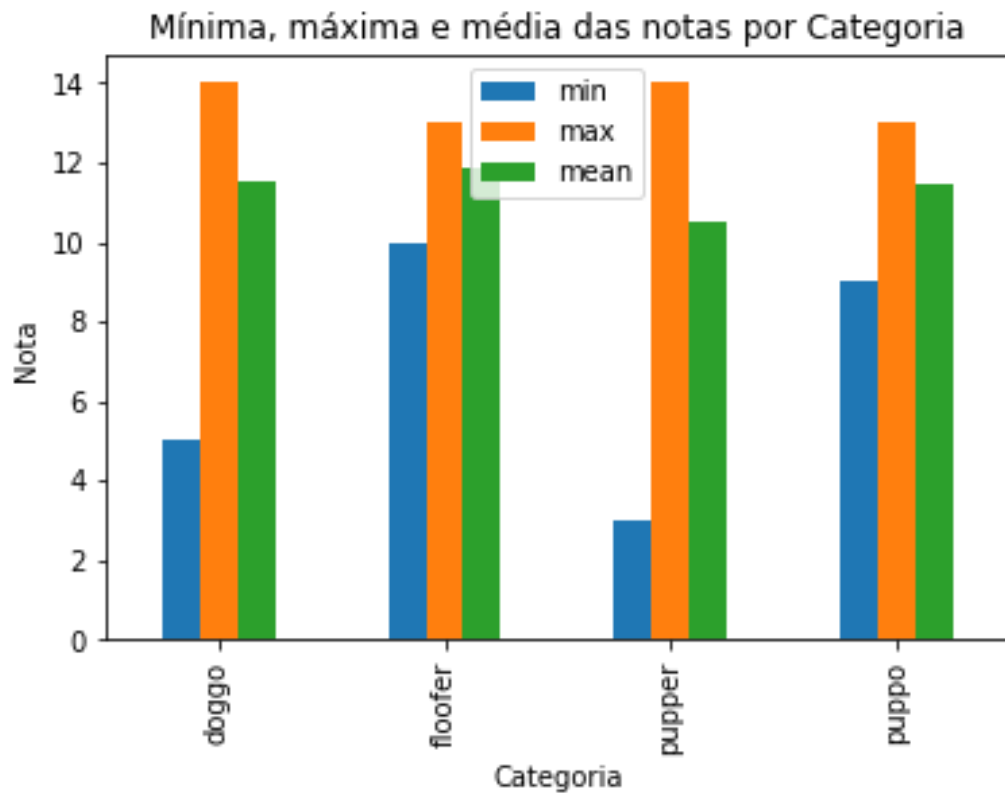
As categorias são **pupper**, **doggo**, **puppo** e **floover** e seu significado pode ser encontrado no *dogtitionary* elaborado pela equipe do @WeRateDog:



Assim, dentre os **257** registros classificados, **188** são classificados como **pupper**, que em tradução livre significa "1. Um pequeno cão, normalmente jovem. Podem ser equivalentes ou menos maduros que alguns doggos. 2. Um doggo inexperiente, não familiar, ou despreparado para responsabilidades atribuídas a um doggo".



3. Quais foram as notas por Categoria? (mínima, máxima e média)

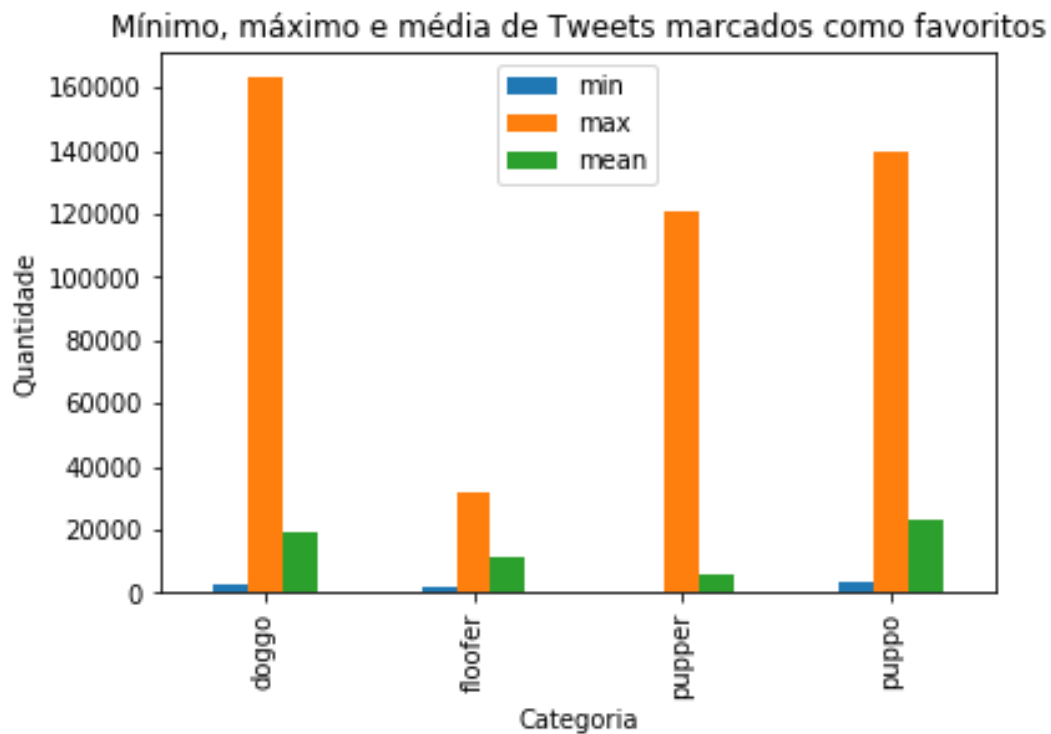


Legenda:

Categoria	Mínima	Máxima	Média
doggo	5	14	11.52
floofer	10	13	11.83
pupper	3	14	10.48
puppo	9	13	11.46

Sobre o quadro de avaliações por categoria temos um empate entre **doggo** e **pupper** na nota máxima, **14**. A categoria **pupper** apresenta a menor nota, **3**. Entretanto, mesmo sem se destacar entre a máxima e a mínima, a categoria **floofer** obteve a melhor média, **11,83**.

4. Qual foi a contagem de Favoritos por Categoria? (mínima, máxima e média)

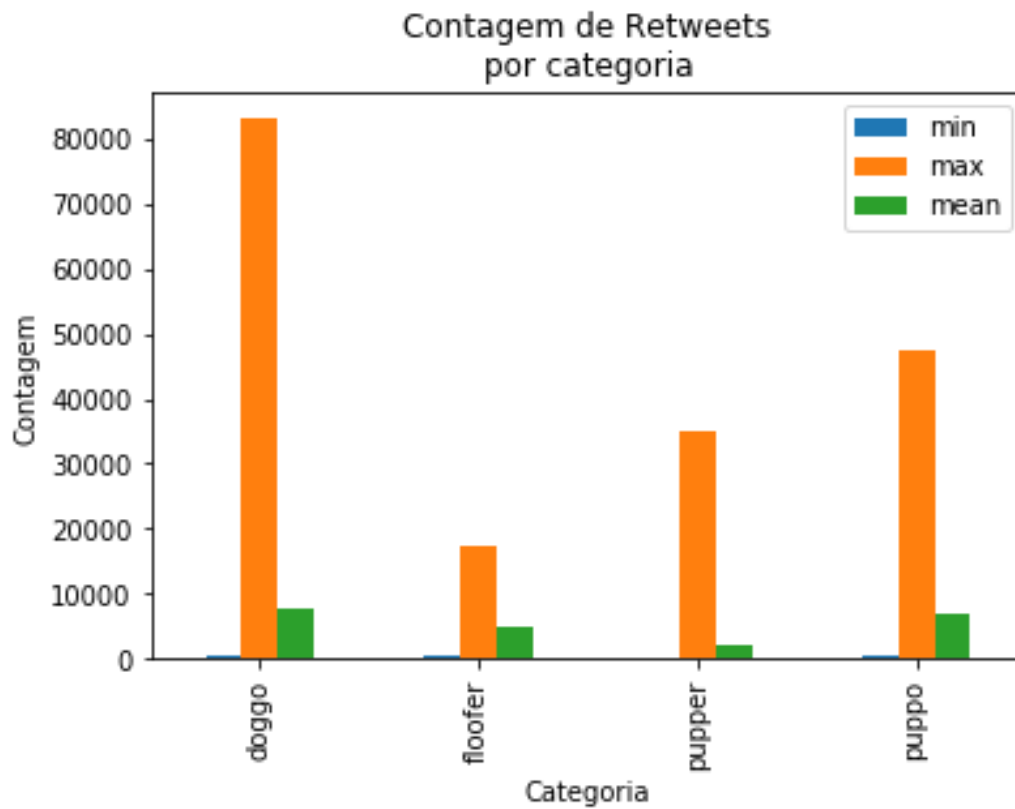


Legenda:

Categoria	Mínima	Máxima	Média
doggo	2458	163.034	19.408,92
floofer	2152	31.685	10.981,66
pupper	661	121.093	6033,87
puppo	3093	139.348	22.861,53

A categoria **doggo** apresenta os Tweets com maior quantidade de favoritos com o máximo em **163.034**.

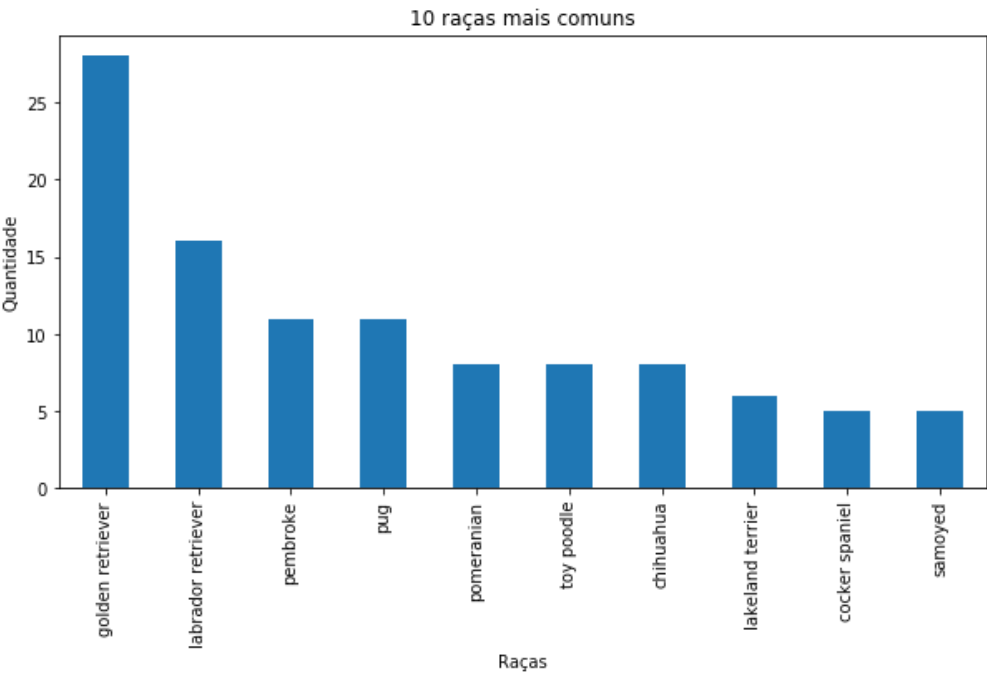
5. Qual foi a contagem de Retweets por Categoria? (mínima, máxima e média)



Legenda:

Categoria	Mínima	Máxima	Média
doggo	675	82.905	7.675,86
floofer	465	17.234	4.813,83
pupper	96	35.138	2.004,73
puppo	658	47.253	7.113,76

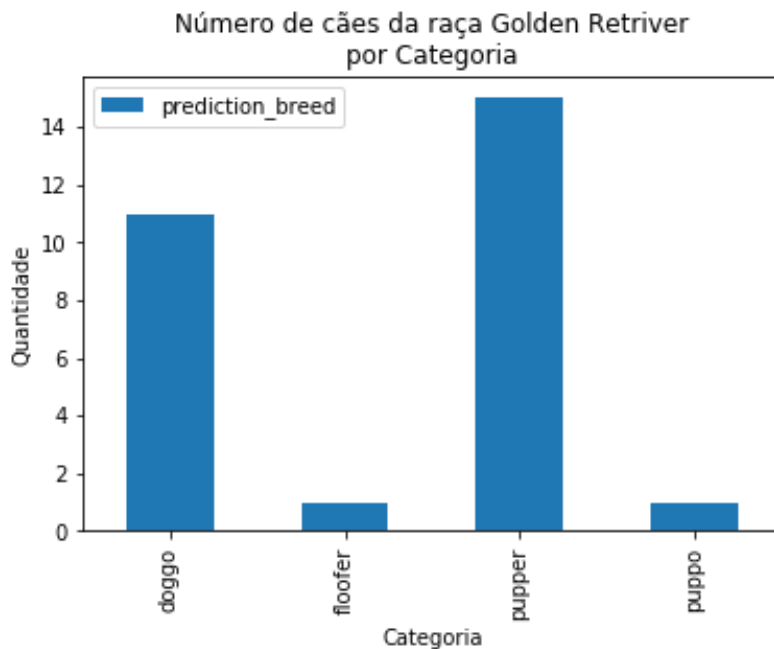
6. Quais são as 10 raças mais populares?



Legenda:

Raça	Quantidade
golden retriever	28
labrador retriever	16
pembroke	11
pug	11
pomeranian	8
toy poodle	8
chihuahua	8
lakeland terrier	6
cocker spaniel	5
samoyed	5

7. Da raça mais popular, qual é a contagem por Categoria?



Legenda:

Categoria	Quantidade
doggo	11
floofer	1
pupper	15
puppo	1

Conclusão

Este projeto teve como foco o processo de coleta, limpeza e união dos dados, assim, não tem propósito conclusivo e cabe perguntas e novas combinações sobre os dados disponíveis.