Autor: Márcio Souza de Oliveira

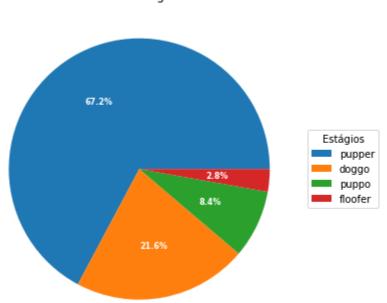
E-mail: emaildomso@gmail.com

## Relatório de Insights

## 1) Estágios de um cão

Apesar da grande quantidade de tweets realizados, apenas uma pequena parte é classificada pelo autor. Nessa amostra temos **1.686** tweets originais (que não são retweets) onde apenas **250** tweets receberam uma classificação (Doggo, Puppo, Floofer, Puppo)

E podemos notar que a grande maioria dos tweets avaliados correspondem a "**Pupper**". Talvez o autor tenha uma grande preferência por filhotes (E quem não tem, rs?)



% de Estágios

Também é notável que a classificação faz alguma diferença quando olhamos a quantidade média de retweets e favoritos em comparação aos tweets não classificados (com exceção dos "Puppers" que curiosamente foram os mais classificados pelo autor, como já vimos)

|                   | Favoritos    | Retweets    |
|-------------------|--------------|-------------|
| puppo             | 22204.142857 | 6392.952381 |
| doggo             | 19130.000000 | 6870.851852 |
| floofer           | 12947.714286 | 4823.571429 |
| não classificados | 8806.404596  | 2599.471448 |
| pupper            | 7480.892857  | 2373.505952 |

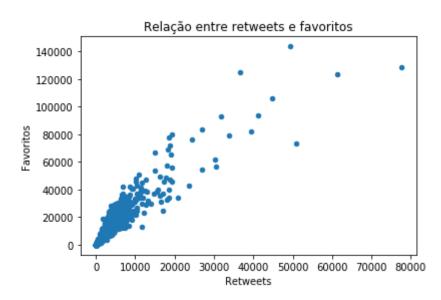
## 2) Avaliações

Podemos observar que a avaliação do autor pesa no sucesso de um tweet. A maior parte das avaliações feitas possuem um numerador entre 10 e 12, com um valor mediano de 1.394 retweets e 4.347 favoritos. Já os tweets com avaliações maiores que 12 possuem um incrível valor mediano de 4.031 retweets e 17.148 favoritos enquanto os que tem avaliações menores de 10 possuem um valor mediano de apenas 587 retweets e 1.763 favoritos.

| Numerador (estatisticas) |             |  |
|--------------------------|-------------|--|
| count                    | 1686.000000 |  |
| mean                     | 11.465599   |  |
| std                      | 8.045844    |  |
| min                      | 0.000000    |  |
| 25%                      | 10.000000   |  |
| 50%                      | 11.000000   |  |
| 75%                      | 12.000000   |  |
| max                      | 165.000000  |  |

|                     | 10~12  | < 10   | > 12    |
|---------------------|--------|--------|---------|
| Retweets (mediana)  | 1394.0 | 587.5  | 4031.0  |
| Favoritos (mediana) | 4347.5 | 1763.0 | 17148.0 |

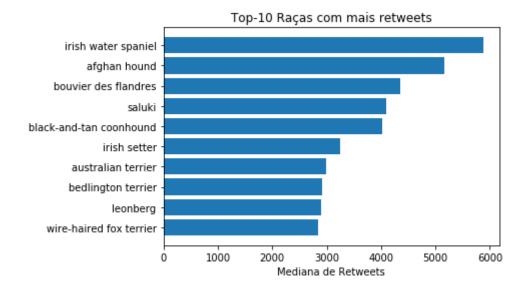
É possível constatar uma forte relação entre retweets e favoritos.



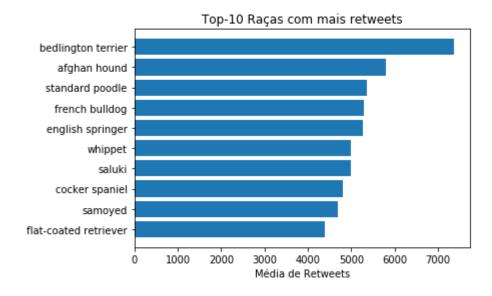
Com isso em mente, continuarei focando minha análise apenas nos retweets como indicador de popularidade de um tweet.

## 3) Raças

Como última análise, podemos obeservar as top-10 raças mais populares



Quando montamos o ranking pelas médias, vemos uma coisa curiosa: a raça **"Bedlington Terrier"** dispara no ranking. Seria algum *outlier* ou erro não tratado no conjunto de dados?



Olhando os tweets da raça Bedlington Terrier percebemos um em especial com uma grande quantidade de retweets (30.302) e favoritos (61.899). Fazendo uma análise mais apurada nos dados do tweet, percebemos que não há nenhum erro nos dados:



https://twitter.com/dog\_rates/status/678399652199309312

Afinal, eu mesmo estou "retuitando" essa fofura e não consigo mais parar de assistir esse video...;-)