

Our favorite doggies



Mohamed Abed
15/9/2020
Act Report

Insights from our data:

After assessing and cleaning our data that has been brought by our twitter api, we got to discover quirky and fun relationships between different metrics that may be a causation or correlation relationship and can be investigated more to find out which one is more accurate.

But until then we'll just put some of these insights on display here and also use some visualization to aid us in portraying them.

First insight:

We can notice how on average the number of favorites for the dog breeds are sorted out, we can discover from this the most loveable for the dog owners' community online.

We found out that most loveable dog breeds that showed great number of favorites on average, and the results were these few displayed in the figure, which had the number one spot for Saluki, followed by Bedlington terrier, and flat coated retriever.

```
In [36]: archive_df_clean.groupby('breed_pred')['favorite_count'].mean().sort_values(ascending=False)

Out[36]: breed_pred
Saluki                24060.000000
Bedlington_terrier    21153.166667
flat-coated_retriever 20241.888889
French_bulldog        18416.806452
Bouvier_des_Flandres  18032.000000
...
Tibetan_terrier       1841.000000
Irish_wolfhound        1534.000000
Japanese_spaniel       1362.000000
groenendael            1156.500000
Brabancon_griffon       885.000000
Name: favorite_count, Length: 114, dtype: float64
```

Second Insight:

We tried to find if the twitter account 'WeRateDogs' had a certain loveable way to call their favorites so searched the history of ratings for every type such as 'doggo', 'floofer', 'pupper' and such. To find out if the rating system can be expected to rate a certain number if a certain word is used in the text.

```
In [37]: archive_df_clean.groupby('dog_stage')['rating_numerator'].mean().sort_values(ascending=False)

Out[37]: dog_stage
puppo      12.043478
floofer    11.875000
doggo      11.794521
pupper     10.727273
Name: rating_numerator, dtype: float64
```

Third Insight:

Plotting the top 10 dog names that got retweeted the most, showing the retweets number for each name to see if some names had a better luck finding their way into twitters users' hearts.

Seeing how that might have been true, we found the results to have a lot of favorite dog names that get more retweets than usual for the average dog, starting with the two dominant names 'Stephan' and 'Duddles', which showed the great love and attention we might show for a dog with a similar name.

Visualization

Most Retweeted Dog Names

