Wrangle and Analyze Data Report

Author: Khawla Alqarni
Date: Dec 12, 2022
Purpose: document data finding and insights.

Who is WeRateDogs? What purpose from analyzing this account's tweets?

WeRateDogs (@dog_rates) is a very popular Twitter account. It has 9.733M followers. Each twitter has a image and its unique rating system. In this project, we have three kinds of datasets, Twitter archive data including the basic tweet information, image prediction that have done by a neural network and additional data extracted from Twitter API.

After looking at the datasets, I've been interested in finding that which dog stage is the most popular and which dog breed is the most popular, which device that most used and the most popular dog names.

Which tweet has the most retweets?

Tweet for 'Labrador_retriever' breed from 'doggo' class have largest number of retweet !! 79515

```
master[master.retweet_count ==master.retweet_count.max()]
```

retweet_count	favorite_count	jpg_url	img_num	breed	conf
79515.0	131075.0	https://pbs.twimg.com/ext_tw_video_thumb/744234667679821824/pu/img/1GaWmtJtdqzZV7jy.jpg	1.0	Labrador_retriever	0.825333

Which tweet has the most favorites?

Tweet for Lakeland_terrier breed from puppo class have largest number of Favorite !! 132810
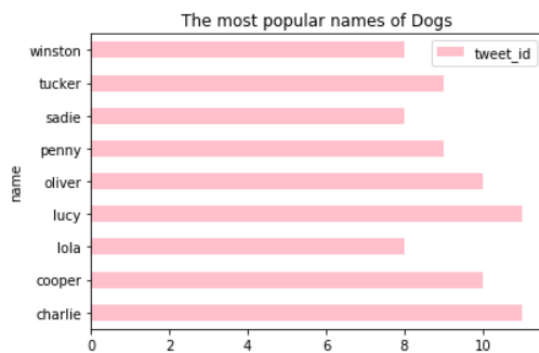
```
master[master.favorite_count == master.favorite_count.max()]
```

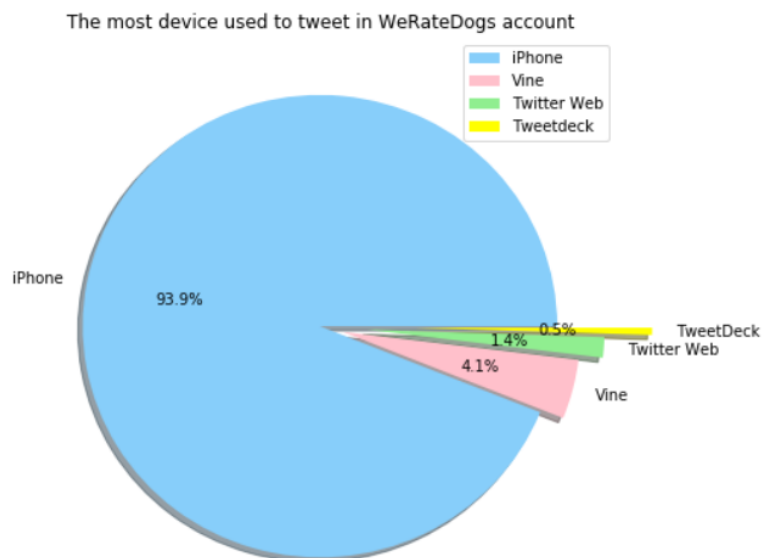| retweet_count	favorite_count	jpg_url	img_num	breed	conf	timestamp	source	text	nam
48265.0	132810.0	https://pbs.twimg.com/media/C2tugXLXgAArJO4.jpg	1.0	Lakeland_terrier	0.196015	2017-01-21 18:26:02	iPhone	Here's a super supportive puppo participating in the Toronto #WomensMarch today. 13/10 https://...	Na

What is the most popular dog's names?

In the figure, Charlie and lucy almost same , they are the most popular names in addition to Cooper and Oliver

```
master.groupby('name').count()[['tweet_id']].query('tweet_id >= 8').plot(kind = 'barh', color='pink')
plt.title('The most popular names of Dogs')
```

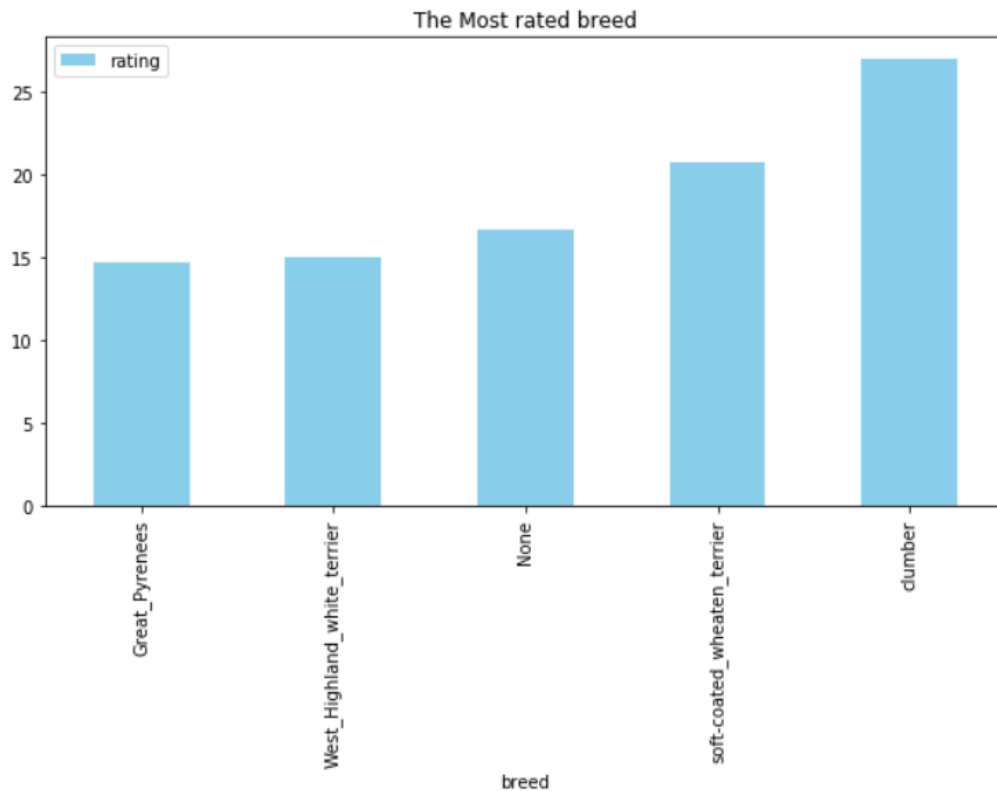]: Text(0.5,1,'The most popular names of Dogs')

What is the most device used to send tweets in this account?

As we can see from pie chart below, iPhone take 93.9% which is a huge percentage among the other devices

What is the most rated breed of dogs?

Clumber is the most rated, we can observe that it has loved by rater!!

The Most rated breed

And finally I wanted to see aggregated values for each breed, count of all retweets and favorites in addition to average rating, we see 'Labrador_retriever' breed is the TOP!!

	breed	retweet_count	favorite_count	rating
1037	Labrador_retriever	79515.0	131075.0	1.3
259	Chihuahua	56625.0	0.0	NaN
533	Chihuahua	56625.0	107015.0	1.3
816	NaN	52360.0	0.0	NaN
1077	Eskimo_dog	52360.0	75163.0	1.3