

# **Report on The Analysis and Visualization of WeRateDogs**

## **Introduction**

This report explains my endeavor to analyze and visualize the WeRateDogs Master Clean dataset. Haven took the time to extensively gather, explore and clean my dataset using three different methods from three different sources, I merged the datasets into a Master clean Dataset, which was saved to file as `twitter_archive_master.csv`. I then visualized the following concerns to gain further insight into the dataset:

- What's the highest source of tweets and ratings?
- Which year has the highest number of Users?
- Which Dog Stage has the Highest Rating?
- What is the most popular Dog breed?
- What are the Ten most popular Dog names?
- What is the correlation between Retweet count and Favorite count?
- Which Tweets ID and Photo has the highest likes?

## **Exploring the Dataset**

In an attempt to analyze the concerns and visualize my finding on the concerns I had earlier put forth, I took my time to think about the best visualization approach and method that best satisfies each concern. Once I have decided on the approach and method, I set forth to explore the dataset below.

However, before a gainful exploration for impactful insight and visualization, the master clean dataset needs to be read to memory. This was done using the usual pandas `'pd.read_csv()'` function.

## Visuals from Insight

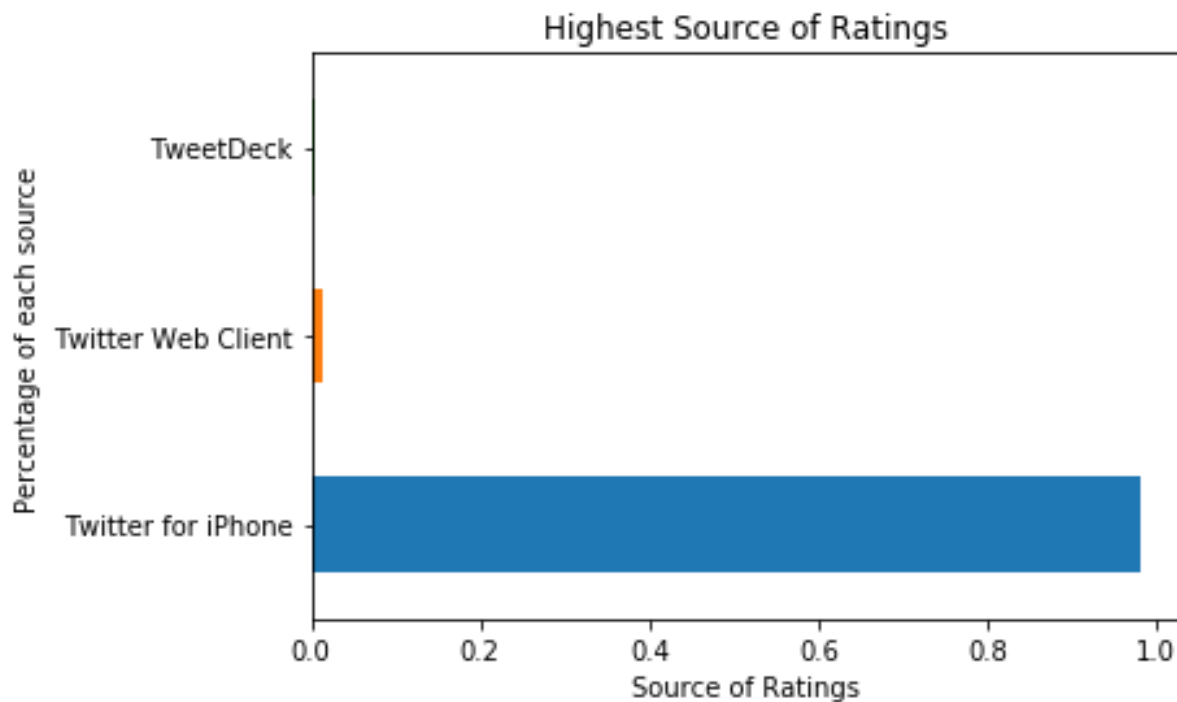
As earlier stated, the following insight is sought from the dataset:

### - What's the highest source of tweets and ratings?

It was discovered that there are three sources from which users tweeted and rate dogs on this platform. The sources and their collective percentage are:

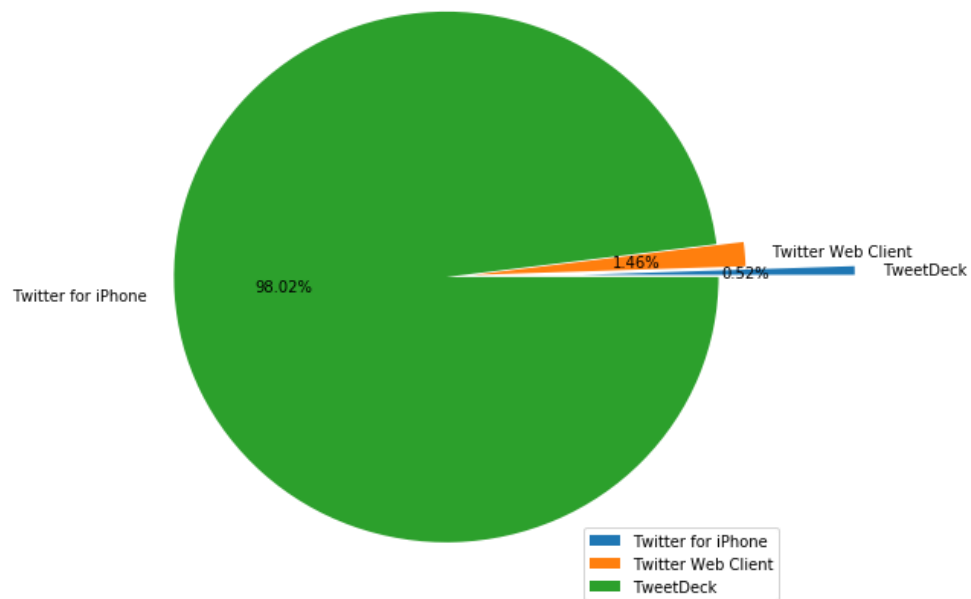
- Twitter for iPhone 98.02%
- Twitter Web Client 1.46%
- TweetDeck 0.52%

The same insight is represented visually, below.



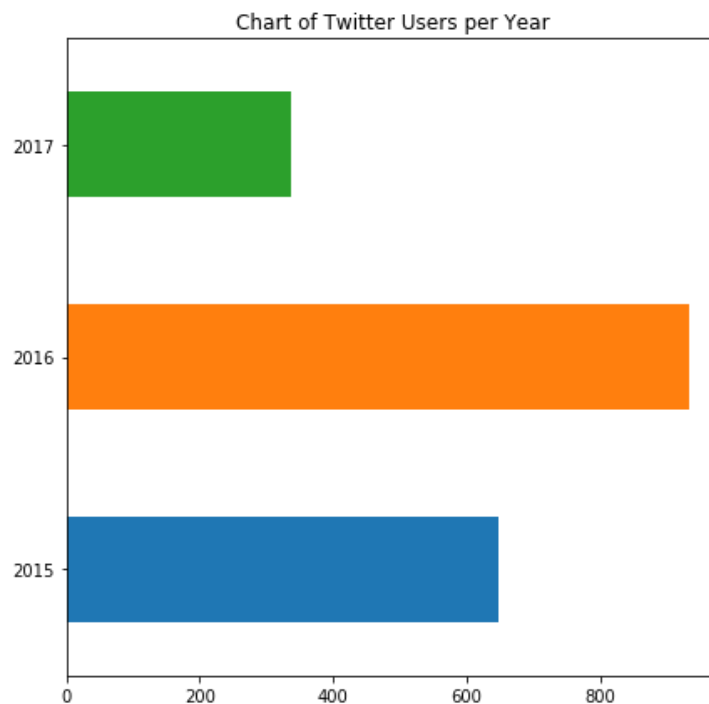
Should the managers of WeRateDogs decide to create a mobile product, they should most definitely focus more on this niche. The same plot is visualised as a pie chart below.

Source of Tweets



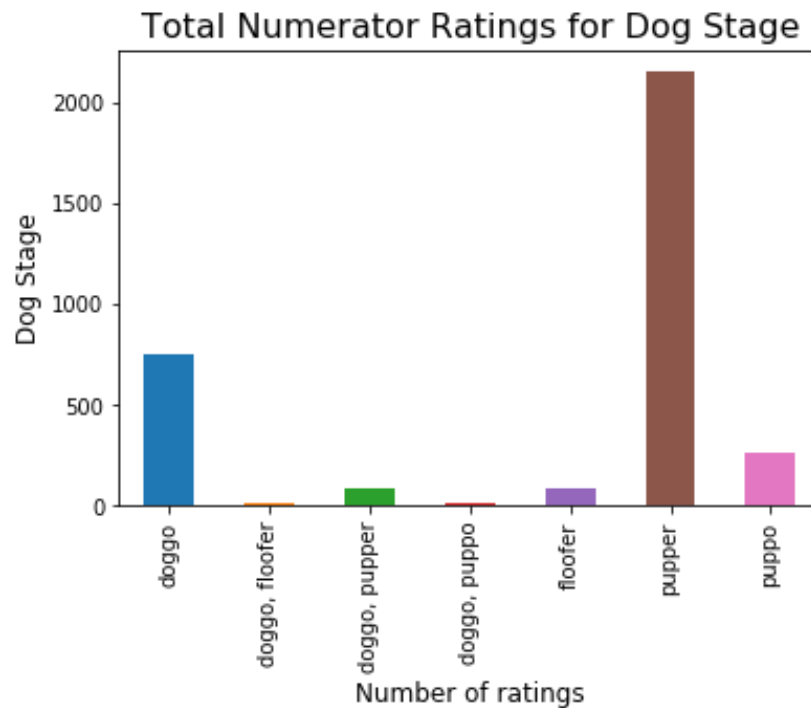
- Which year has the highest number of Users?

Results from the analysis show that in 2015, the platform had 647 users, in 2016, the platform had 932 users and in 2017, it had 338 users.

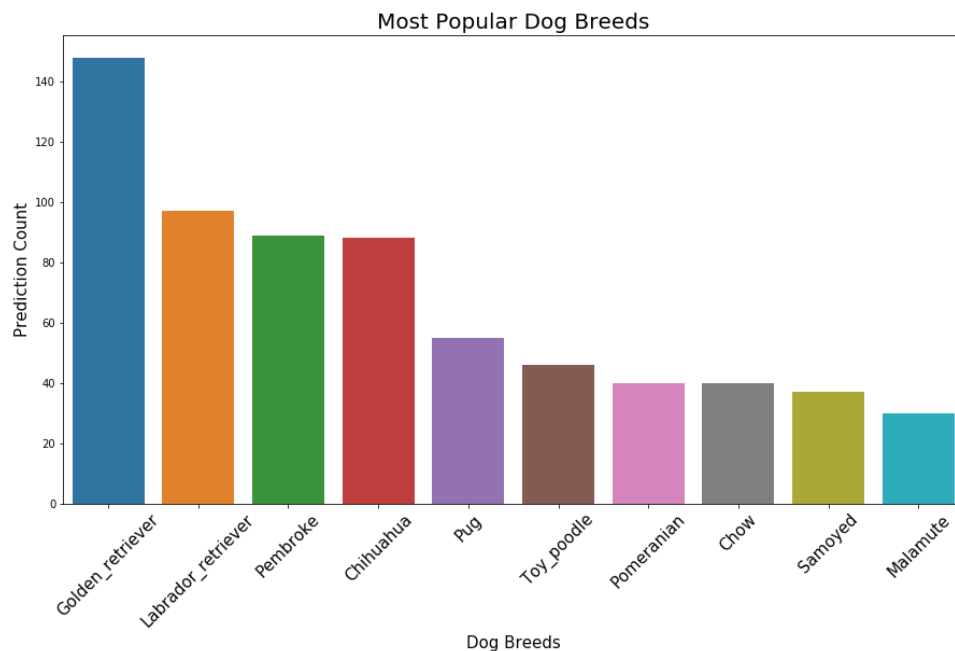


- Which Dog Stage has the Highest Rating?

This bar chart clearly shows that pupper dog stage, has the highest number of ratings, while the doggo-floofer dog stage has the least ratings.



- What is the most popular Dog breed?

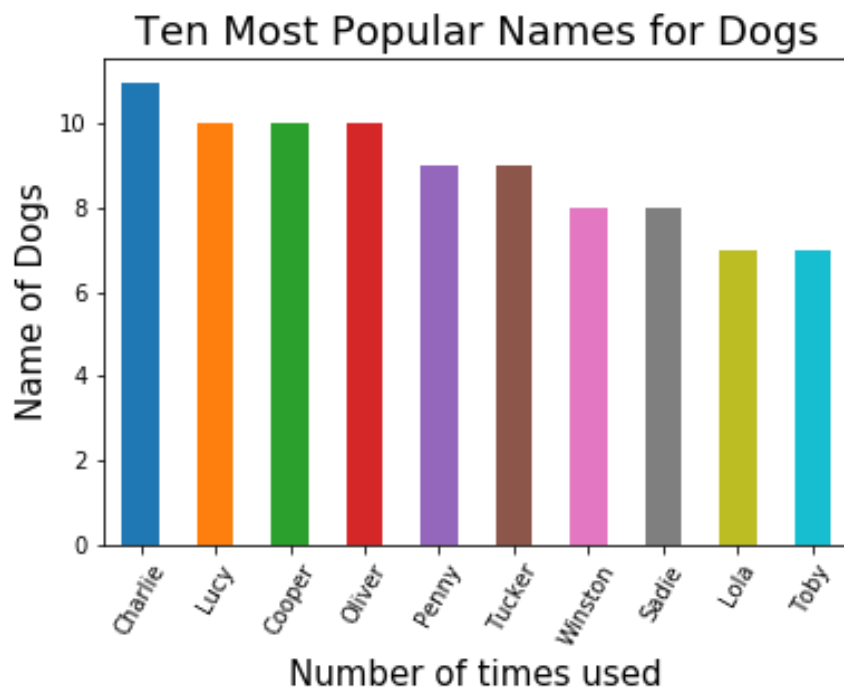


Golden Retrievers seemly appear to be the most popular dog breed on WeRateDogs platform, followed by Labrador Retriever and Pembroke, respectively.

- **What are the Ten most popular Dog names?**

The ten most popular dog names are given below in descending order.

Charlie	11
Cooper	10
Lucy	10
Oliver	10
Tucker	9
Penny	9
Sadie	8
Winston	8
Toby	7
Lola	7

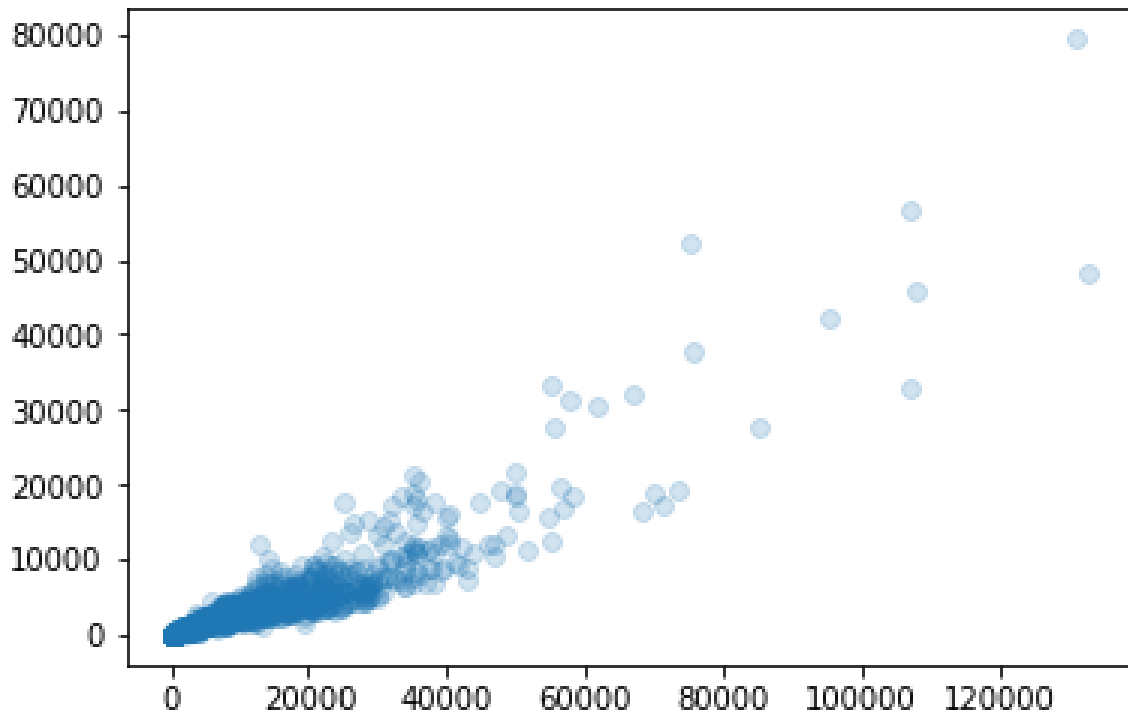


- **What is the correlation between Retweet count and Favorite count?**

Careful analysis show that Retweet count and Favorite counts have a correlation equal to:

0.91233570262597619

This means that has retweet count increase, so also does favorite counts and vice versa. The scatter plot is shown below.



- **Which Tweets ID and Photo has the highest likes?**

	tweet_id	favorite_count	jpg_url	dog_stages	dog_breed
1624	822872901745569793	132810	<a href="https://pbs.twimg.com/media/C2tugXLXgAArJO4.jpg">https://pbs.twimg.com/media/C2tugXLXgAArJO4.jpg</a>	puppo	Lakeland_terrier
1174	744234799360020481	131075	<a href="https://pbs.twimg.com/ext_tw_video_thumb/74423...">https://pbs.twimg.com/ext_tw_video_thumb/74423...</a>	doggo	Labrador_retriever
1860	879415818425184262	107956	<a href="https://pbs.twimg.com/ext_tw_video_thumb/87941...">https://pbs.twimg.com/ext_tw_video_thumb/87941...</a>	NaN	English_springer
1537	807106840509214720	107015	<a href="https://pbs.twimg.com/ext_tw_video_thumb/80710...">https://pbs.twimg.com/ext_tw_video_thumb/80710...</a>	NaN	Chihuahua
1813	866450705531457537	106827	<a href="https://pbs.twimg.com/media/DAZAUFbXcAAG_Nn.jpg">https://pbs.twimg.com/media/DAZAUFbXcAAG_Nn.jpg</a>	pupper	French_bulldog



Isn't this dog beautiful? This dog is the most favored by users of WerateDogs

## **Conclusion**

Analyzing and visualising the insights from a dataset help to bring the dataset to life and spur engagement.