# Analyzing and Visualizing Data

In this Data Wrangling project, we analyzed tweets from Twitter user @dog_rates, also known as WeRateDogs. I gathered, assessed, cleaned and stored the project data. After which, I carried out Data Analyzing and Data Visualizing of the cleaned merged master dataset. The master dataset is made up of 3 unique datasets:

1. Enhanced Twitter Archive (Twitter-archive-enhanced.csv): This is a list of tweets of interest provided by Udacity.
2. The tweet image predictions (Image-predictions.tsv: The file is hosted on Udacity's server, it is a neural network output predicting dog breeds from the image provided in the tweet.
3. Data from the Twitter API (Tweet_json.txt): This is a JSON text file that was generated using Twitter API (tweepy) querying the list of tweets from twitter-archive-enhanced.csv.

After analyzing and visualizing the master dataset, these are my insights:

1. Most of the twitter posts were created using Twitter for iPhone.
2. The most common dog breed in the predictions, is the Golden retriever
3. There is correlation between the Retweets and Favourite Counts(likes), meaning most posts with high dog retweets, generated more favourite counts(likes).

VISUALIZATION:

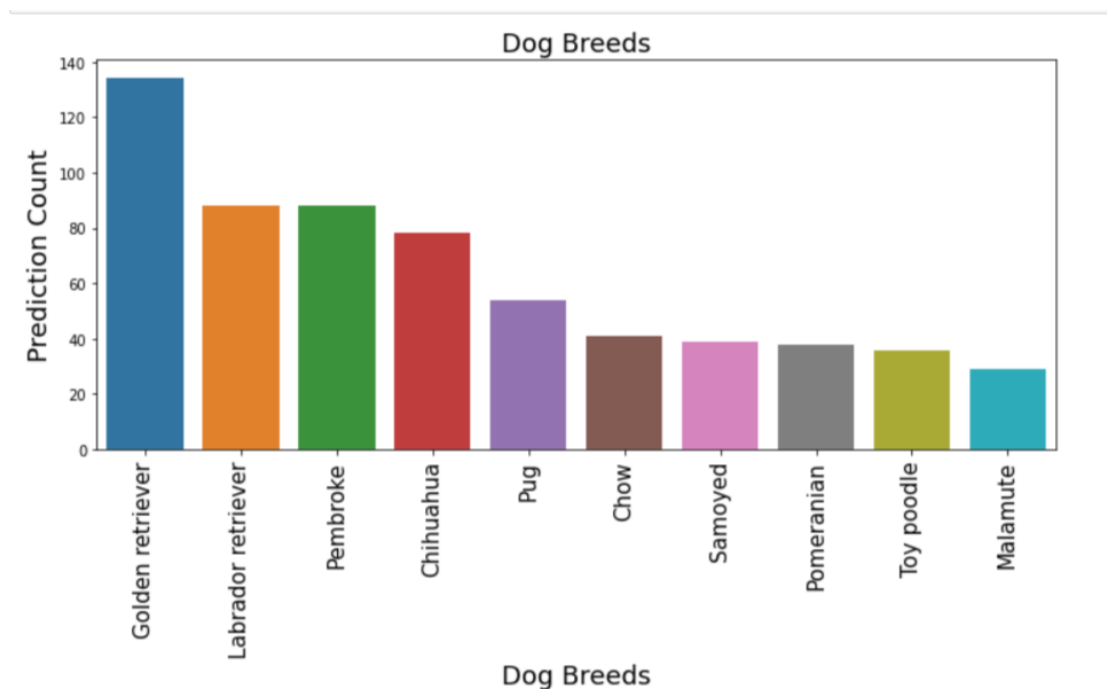1. The species of dogs that most frequently appear on WeRateDogs Twitter account.



Fig. 1. Top 10 most frequent dogs on WeRateDogs Twitter account

The above diagram depicts that the Golden retriver is the most popular breed in the predictions and has the most tweets, with 134 predictions. It is followed by the Labrador retriever and Pembroke, with both having 88 predictions. This suggests that owners of these dog breeds may use twitter more and send dog pictures to WeRateDogs more often than owners of other dog breeds.

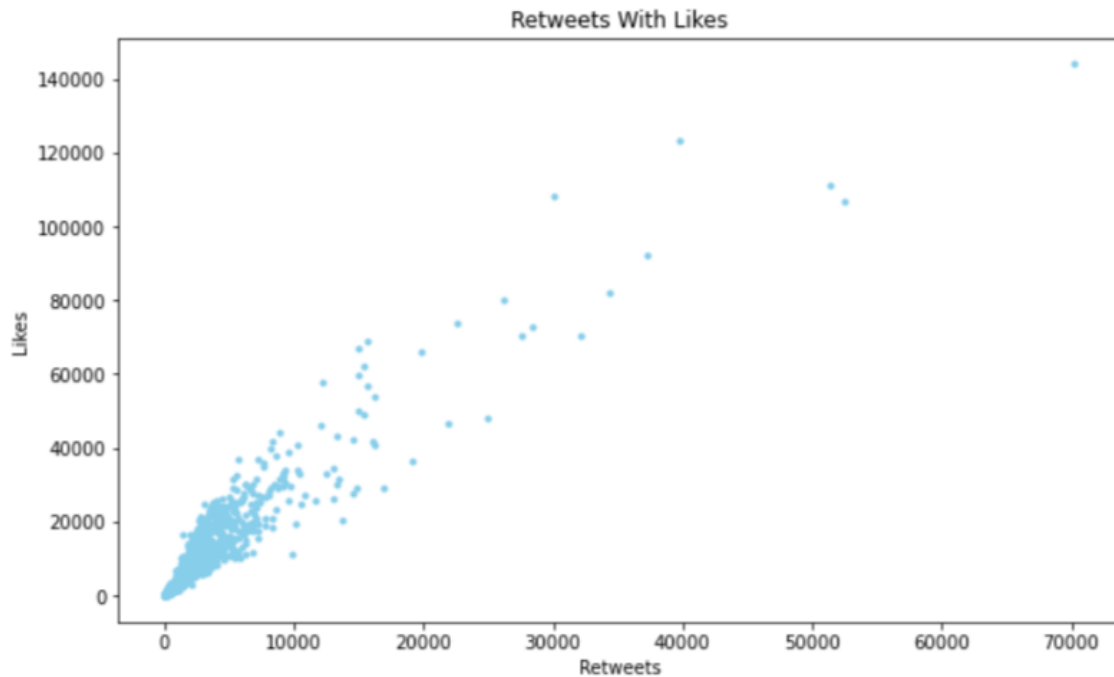2. Correlation between Retweets and Favourites (Likes)



Fig 2. correlation between the Retweets and Favourite Counts(likes)

From the above diagram, there is a correlation between the Retweets and Favourite Counts(likes), meaning most posts with high dog retweets, generated more favourite counts(likes).