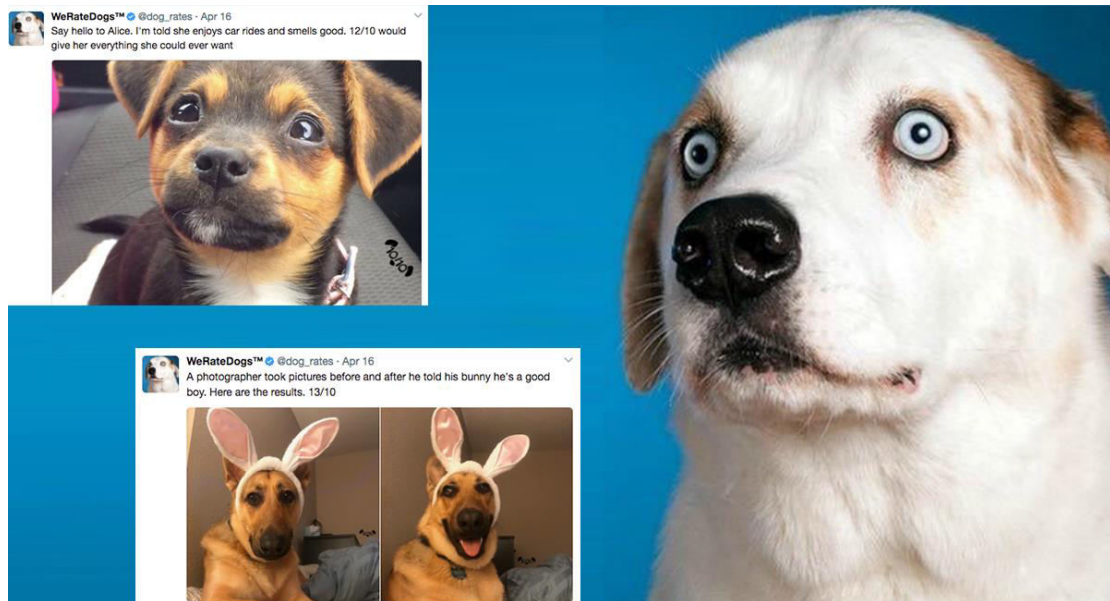


# ANALYSIS AND VISUALIZATION OF THE 'WE RATE DOGS' DATA

BY MUTHOLIB YUSIRA

## INTRODUCTION

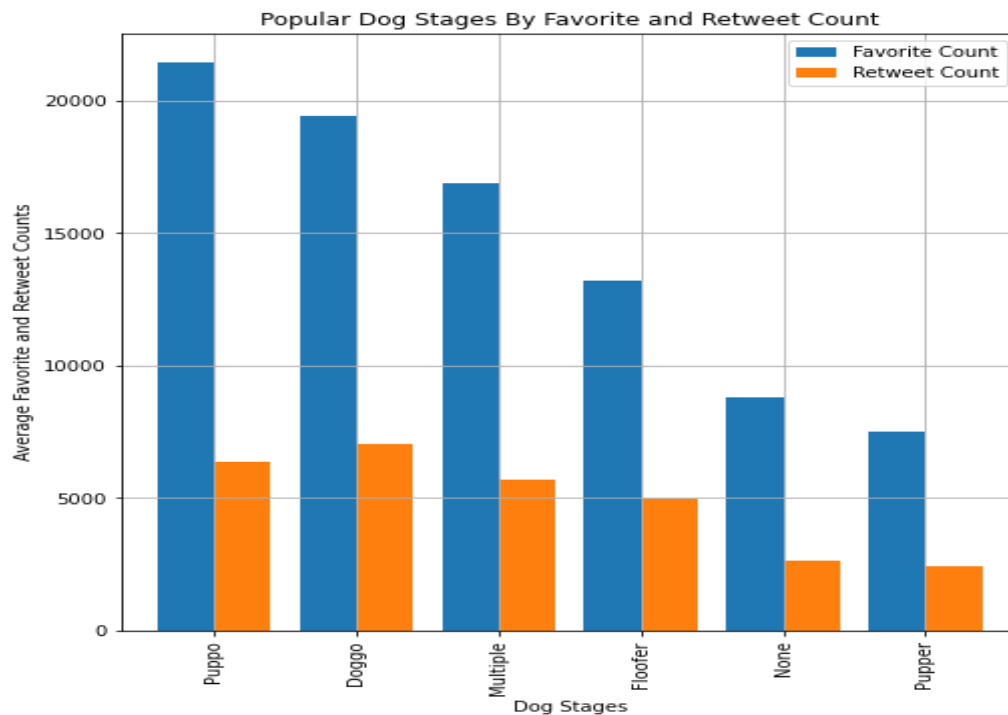
The purpose of this project is to put in practice what I learned in data wrangling data section from Udacity Data Analysis Nanodegree program. The dataset that is wrangled is the tweet archive of Twitter user @dog\_rates, also known as WeRateDogs. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. The goal is to wrangle WeRateDogs Twitter data to create interesting and trustworthy analyses and visualizations. The Twitter archive is great, but it only contains very basic tweet information. Additional gathering from other sources, then assessing and cleaning was done for "Wow!"-worthy analyses and visualizations.



## INSIGHTS AND VISUALIZATIONS

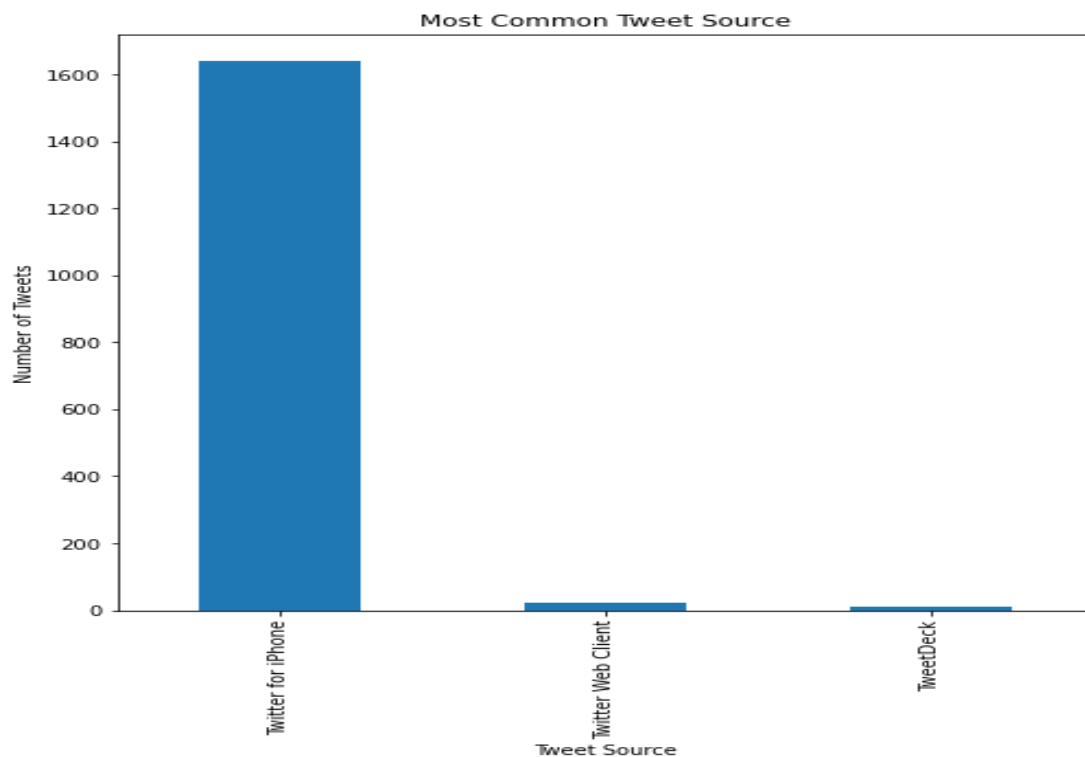
### Insight 1: Most popular dog stage

There are distinct 4 dog stages in the dataset with 2 other categories, none, for those tweets without a dog stage, and multiple, for tweets with more than one dog stages. The most popular dog stage is Puppo with the highest favorite count and second highest retweet count



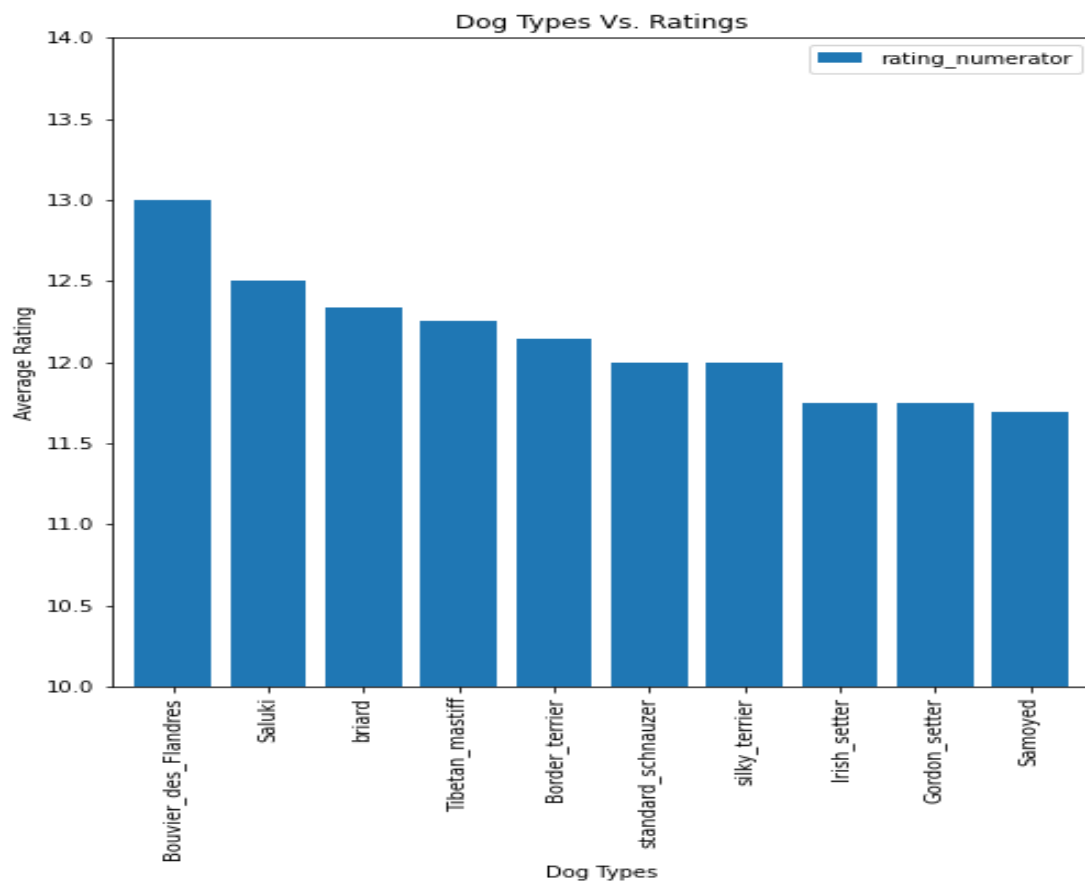
### **Insight 2: Most common source of original tweets**

There are 3 sources of original tweets in the dataset which includes 'Twitter for iPhone', 'Twitter Web Client', and 'TweetDeck'. The most common source of all 3 is the 'Twitter for iPhone' source



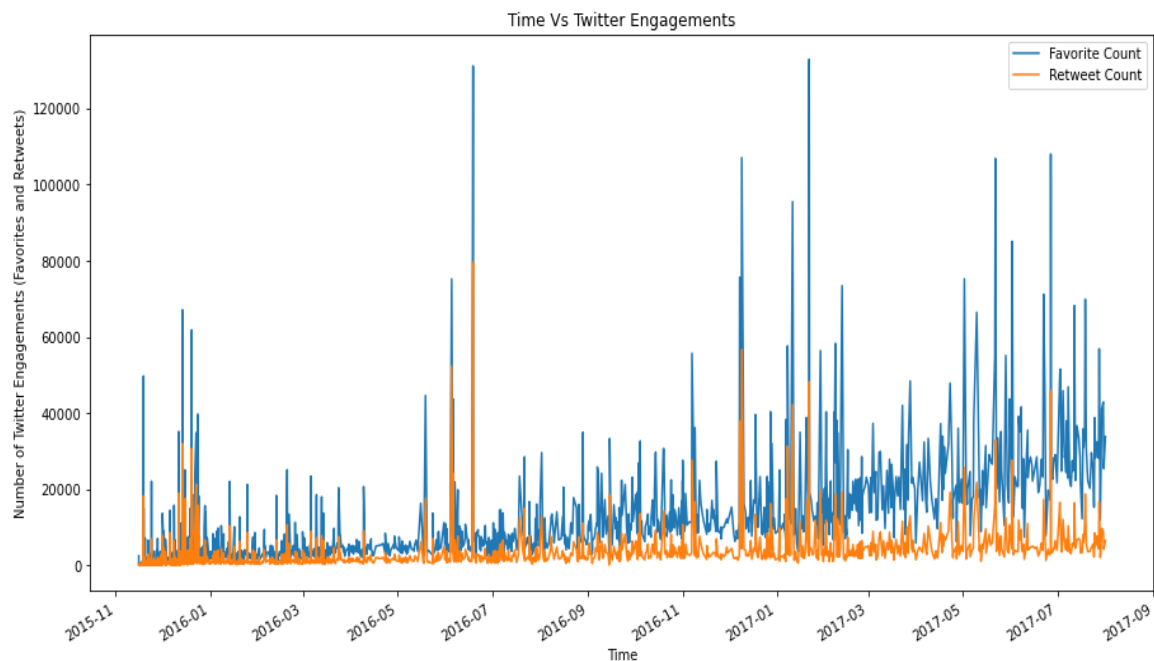
### **Insight 3: Type of dog with the highest rating**

The images in the tweets were predicted to be certain types of dogs. All the dogs had individual ratings. The type of dog with the highest average rating in the dataset is the Bouvier des Flandres dog.



#### **Insight 4: Twitter engagements over time**

The number of twitter engagements with the dog tweets increased over time. There is a more obvious increase in the favorite counts over time than in the retweet counts.



### **Insight 5: Relationship between ratings and the twitter engagements**

There is a positive correlation between the favorite counts and the retweet counts, meaning that the higher the favorite counts, the higher the retweet counts and vice versa. The dogs with the highest ratings also have the highest favorite and retweet counts.

