WE RATE DOGS RETWEETS

❖ INTRODUCTION

- The dataset that we will be wrangling (and analyzing and visualizing) 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. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. Why? Because "they're good dogs Brent." WeRateDogs has over 4 million followers and has received international media coverage.
- WeRateDogs downloaded their Twitter archive and sent it to Udacity to use in this project. This archive contains basic tweet data (tweet ID, timestamp, text, etc.) for all 5000+ of their tweets as they stood on August 1, 2017.

❖ GOAL

- Main tasks in this project are as follows:

  1. Data wrangling, which consists of:
     - Gathering Data
     - Assessing Data
     - Cleaning Data
  2. Storing, analyzing, and visualizing your wrangled data
  3. Reporting on 1) your data wrangling efforts and 2) your data analyses and visualizations

- Sharing tweets is one of the main keys for growing a twitter account.
- As a result, we try to use the data at our disposal towards getting a better understanding on what influences the tweet sharing, such as time (hours, days), kind of dogs (breeds, stages) and rating, in order to give some insight to the marketing team.
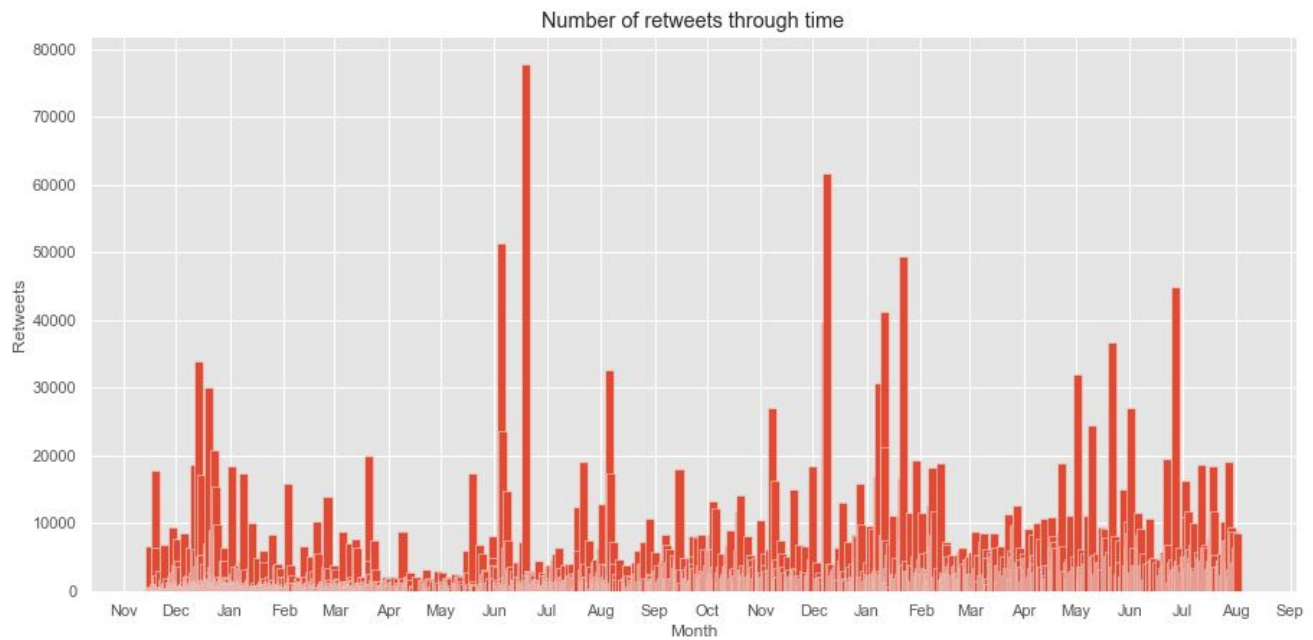
❖ GENERAL

- Let us recall some general statistics regarding the Count and Retweets, prior to starting analysis.

1. Describe:

- ❏ Count : 1971.000000
- ❏ Mean : 2756.662100
- ❏ Standard Deviation : 4749.553787
- ❏ Minimum : 13.000000
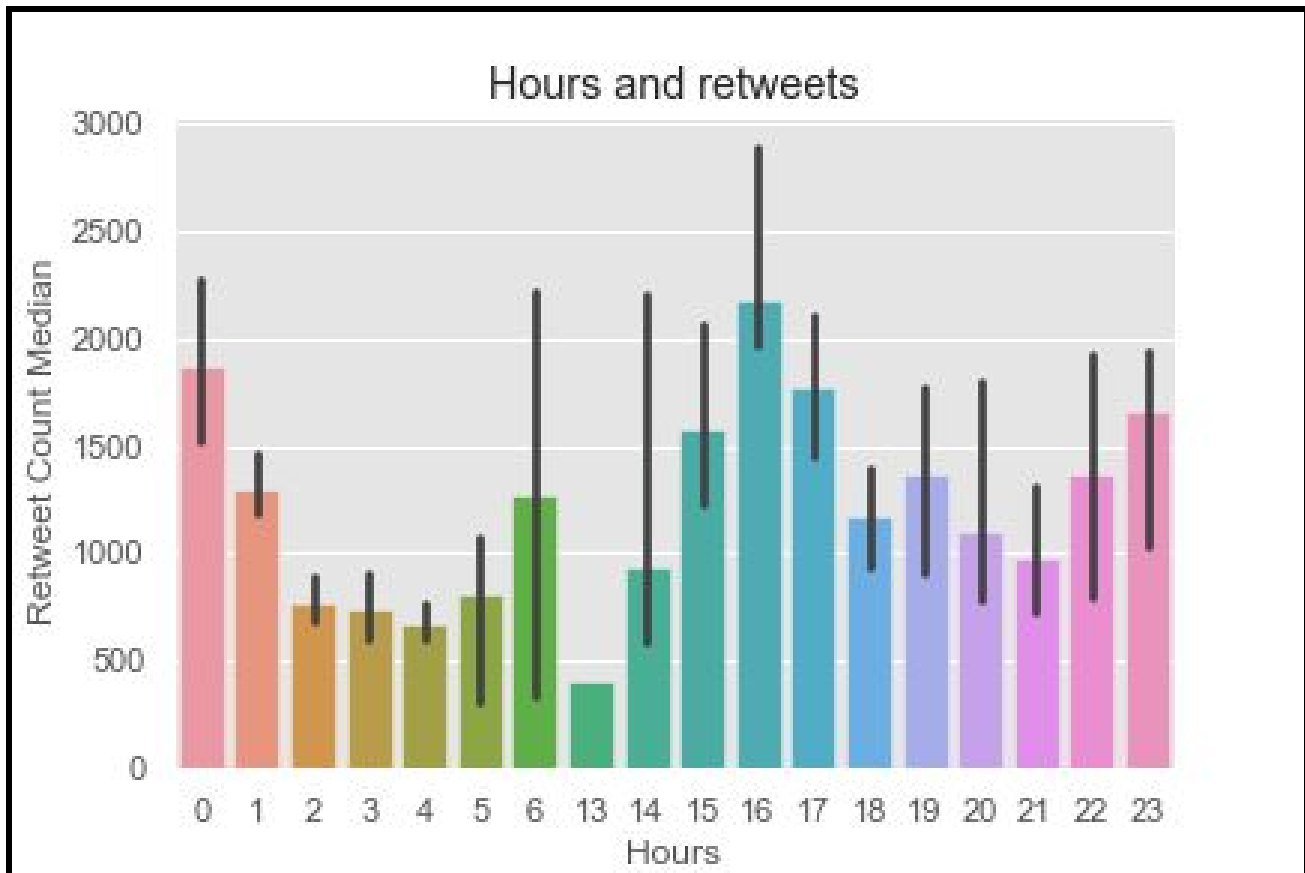- ❏ Median : 3164.000000
- ❏ Maximum : 77897.000000

2. Retweets through TIme:

● Post the removal of Outliers, we perform the specific analysis.
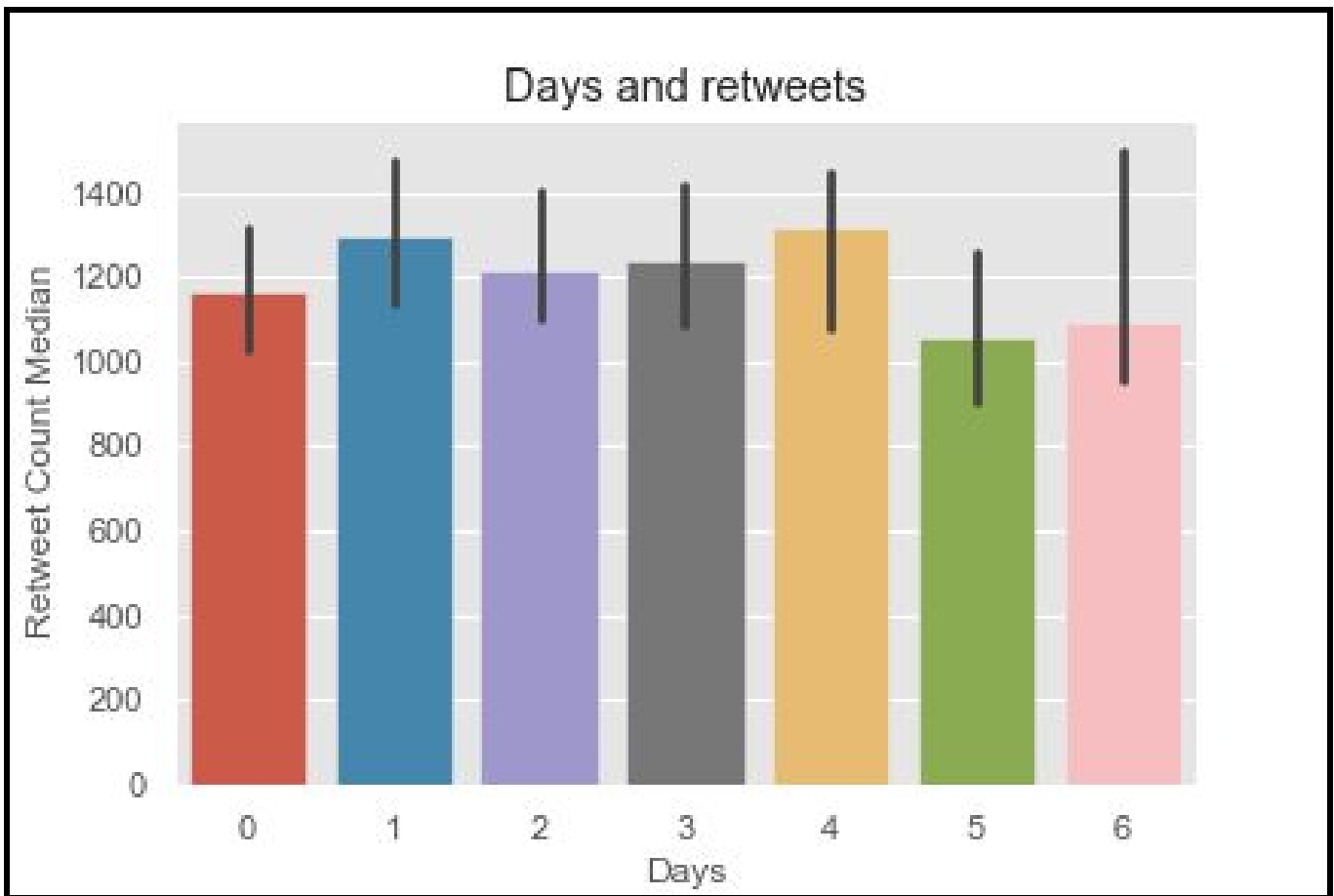
❖ RETWEET Vs TIME

1. Hours Vs Retweets :

- From mere observation we can notice that the count of Retweets is high at some particular timings like 4 PM(16 hours) or 5 PM(17 hours).

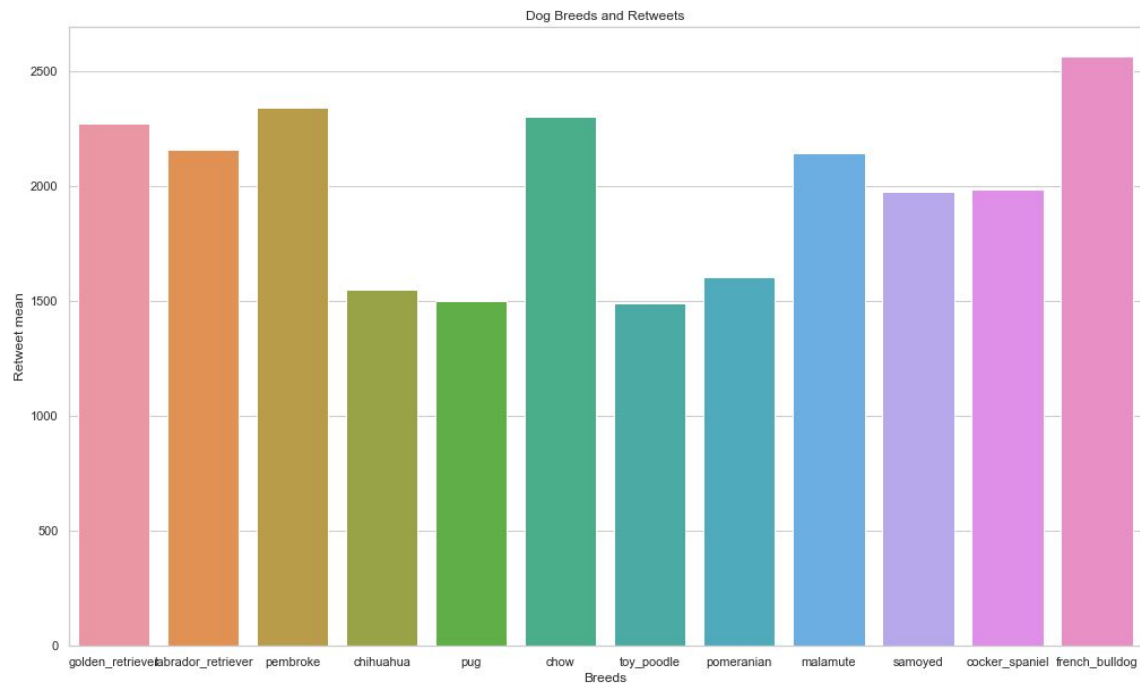- Alternatively, the Retweets show a deep dive in numbers at hours like 3 AM, 4 AM, 1 PM.

2. Days of the Week Vs Retweets :

- Please Note that Monday is represented by 0, Tuesday is represented by 1, ..., and Sunday is represented by 6.

- We can notice that Tuesday and Friday, show a better performance with respect to the Retweets, as compared to Saturday and Sunday.
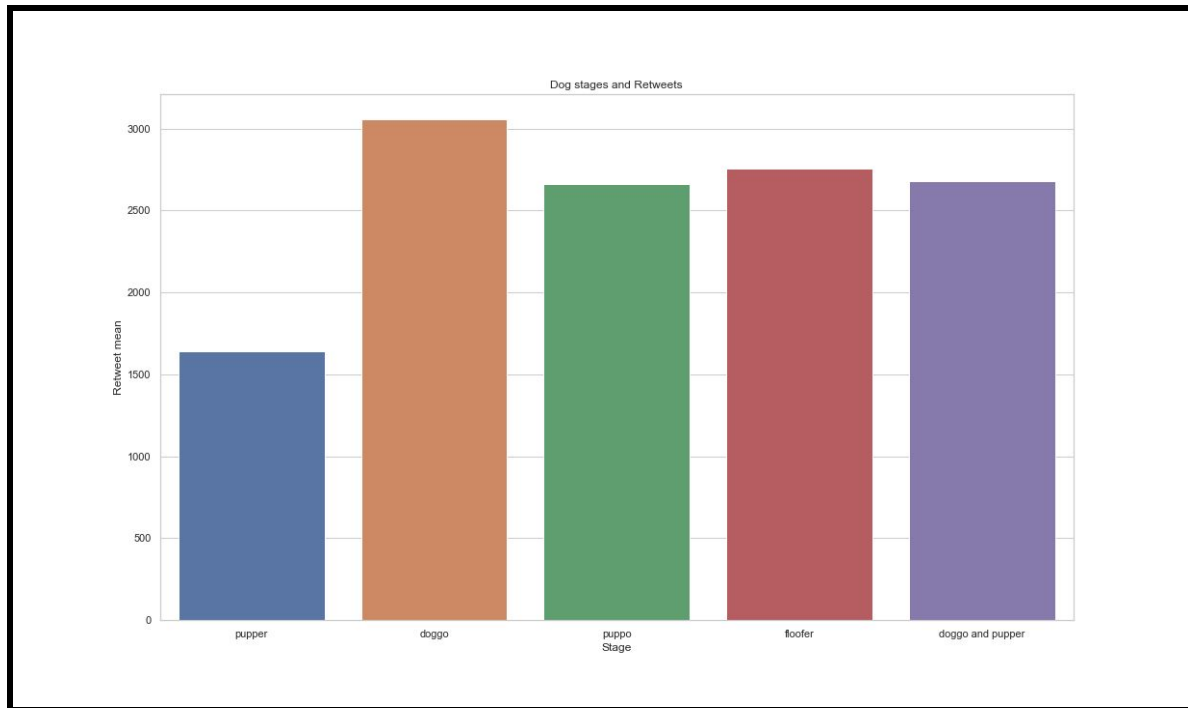
❖RETWEET Vs Kind of Dogs

1. Retweets Vs Dog Breeds

- The above plot depicts the Breeds with the most frequent Retweets (Those having over 23 Retweets over the time frame mentioned in Data Frame.)

- We notice that there are huge differences among different Breeds.

- French Bulldogs have more than 2500 average retweets while pug toy poddles have less than 1500 average retweets.

2. Retweets Vs Dog Stages

- It is pretty evident that the dog stages do not possess similar retweets averages, doggo count has an average superior to 3000 and pupper is only around 1500.