

# Act Report

## Data Insight

Finally, this project got two cleaned and tidy datasets which were saved into two csv files. The one is twitter\_archive\_master.csv, the other one is image\_predictions\_master.csv. After a basic data analysis, I found the following points:

### The tweet which is most retweeted:

“Here's a doggo realizing you can stand in a pool. 13/10 enlightened af (vid by Tina Conrad) <https://t.co/7wE9LTEXC4>”.  
Retweet count = 78616

### The tweet which is least retweeted:

“@bragg6of8 @Andy\_Pace\_ we are still looking for the first 15/10”.  
Retweet count = 0

### The tweet with maximum like:

“Here's a super supportive puppo participating in the Toronto #WomensMarch today. 13/10 <https://t.co/nTz3FtorBe>”  
favorite\_count = 131571

### The tweet with minimum like:

“@serial @MrRoles OH MY GOD I listened to all of season 1 during a single road trip. I love you guys! I can confirm Bernie's 12/10 rating :)”  
favorite\_count = 51

### The mean and median rating:

The mean rating is 12.48 and the median rating is 11. Therefore, the dogs on these pictures are all very good.

### The prediction accuracy:

Prediction1 Accuracy = 73.83%

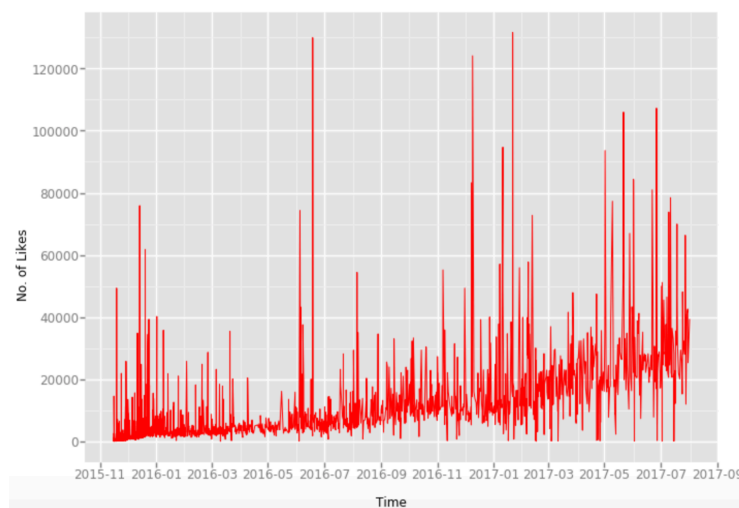
Prediction2 Accuracy = 74.84%

Prediction3 Accuracy = 72.24%

Top three predictions have a high prediction accuracy. The accuracy of prediction 2 is higher than prediction1.

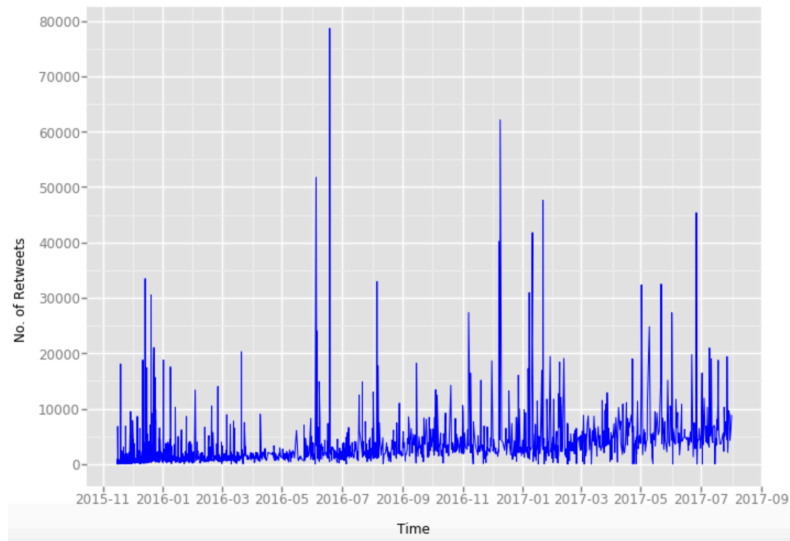
## Data Visualization

### 1. Trends of likes with time:



The number of likes keeps increasing with time and the relation is almost positive linear. Therefore, this twitter is more and more popular with time.

## 2. Trends of retweets with time:



The number of retweets keeps increasing with time and the relation is almost positive linear. Therefore, more and more people pay attention to dogs and follow this twitter.