

# **Data Analysis Report**

This report document my findings from the exploration of data collected from WeRateDog, a twitter account popularly known for rating dogs. WeRateDog often rates dogs above its rating standard of 10. Although some dogs are given rating lower than 10.

In this report I will show how dogs with high rating defer from dogs with low rating. I will find out how the reaction of dog lovers toward dogs with high rated dogs differ from dogs with low rating.

The data will also be explored to know if dog lover sentiment agrees with the poster ratings.

## **Analysis And Visualisation**

There are two groups of dogs explored in the analysis. I referred to one group of dogs as good dogs and the other group as bad dogs. The bad dogs are not really bad. It's only for analysis purpose. I will explain further below.

Groups of Dogs:

- Good dogs in the analysis are dogs that are rated 10 and above.
- Bad dogs are dogs that are given ratings below 10.

This idea was taken from the dog rating creator(@WeRateDogs) when he replied Brant on twitter. Brant had frowned at the poster's attitude of rating dogs over 10. His reply was that such dogs are good dogs.

So, the dog poster determine which dog is good and which is bad.

Please, kindly note that the group referred to as bad dogs are not really bad. it is just a way of describing the physical appearances surrounding the dogs and for the analysis.

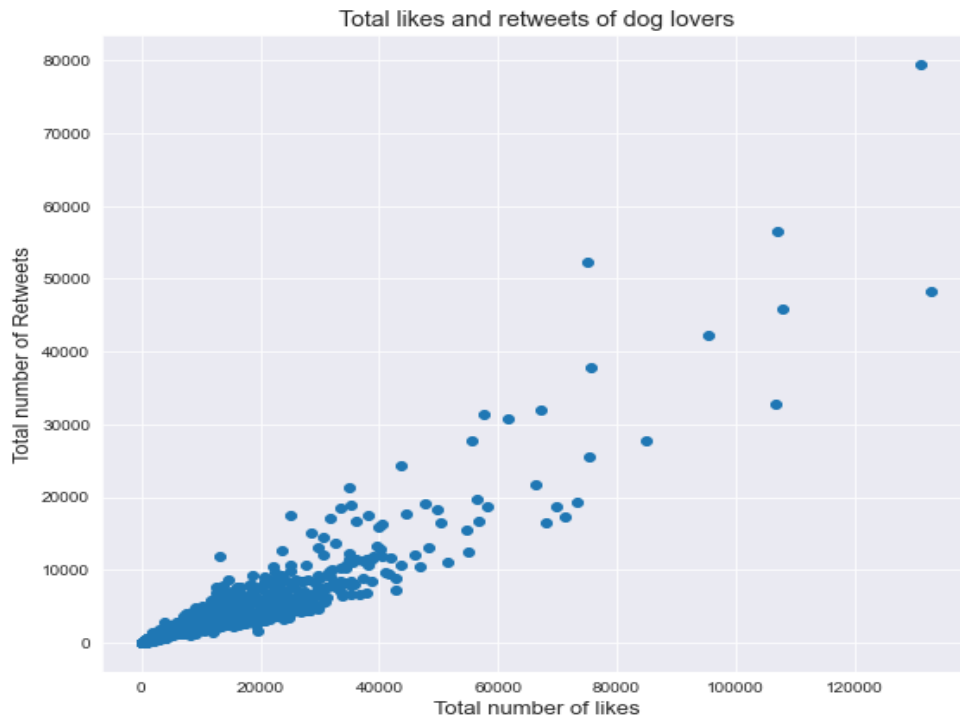
## **Research Questions**

In my exploration, I seek to answer the following question.

1. Did increase in likes increase the number of retweets of dog lovers?
2. Which dogs did dog lovers retweeted more? Good or bad dogs?
3. Which dogs did dog lovers liked more?
4. What is the average rating of both groups of dogs?
5. Which group of dogs has more likes among dog lovers?
6. Which dog group did dog lovers retweeted more?
7. What breed is the dog with the highest number of likes?
8. Which dog breed did dog lovers liked and retweeted the least?

**Did increase in likes increase the number of retweets of dog lovers?**

In answering this question, a scatter plot was plotted to show how the number of likes impacts the number of retweets.



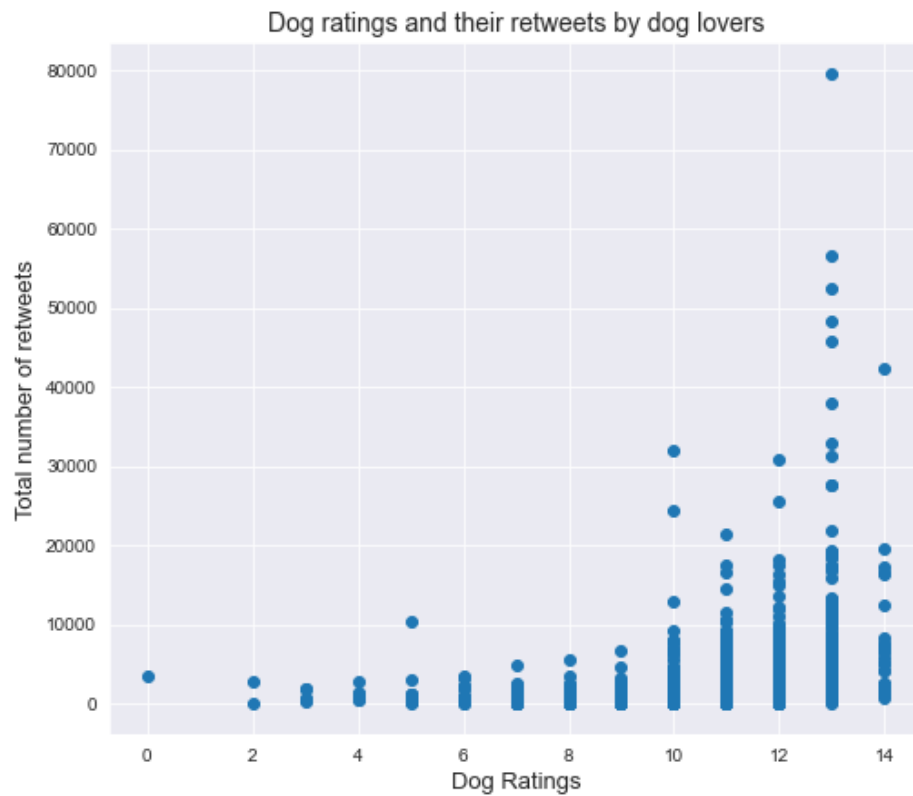
### **Insight:**

I found that there is a positive correlation between the total number of likes and retweets of dog tweets among dog lovers. This means that dog lovers tend to retweet the dog tweets that they liked more. They are more likely to retweet a dog tweet when they like it.

## **Which dogs did dog lovers retweeted more?**

### **Good or bad dogs?**

This question will be answered by plotting the dog ratings and total number of retweets per dog on a scatter plot to see if there is a positive correlation between them.



### Insight:

From the scatter plot above, I found that there is a positive correlation between the dog ratings and total number of retweets among dog lovers.

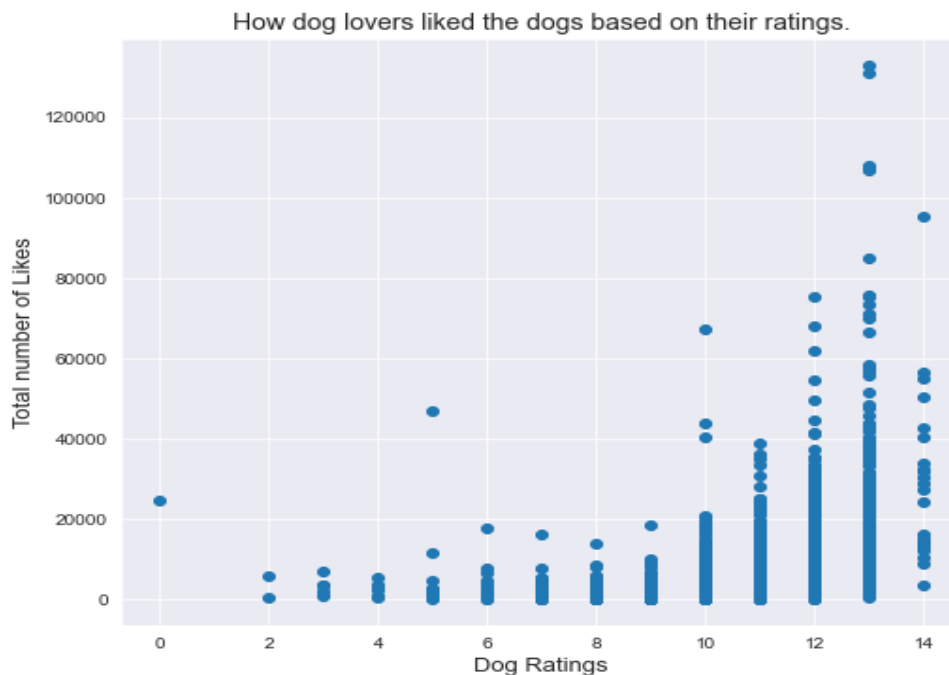
This means that dog lovers tend to retweet posts about good dogs more than they did with bad dogs.

Dog ratings from 10 and above have more retweets than dogs with lesser ratings.

This confirms that the dog poster rating agrees with dog lovers. It can also be deduced from the above visualization that the higher the dog rating, the more the retweets.

## Which dogs did dog lovers like more?

This question will be answered by plotting the ratings of each dog and their total number of likes on a scatter plot to see if there is a positive correlation between them.



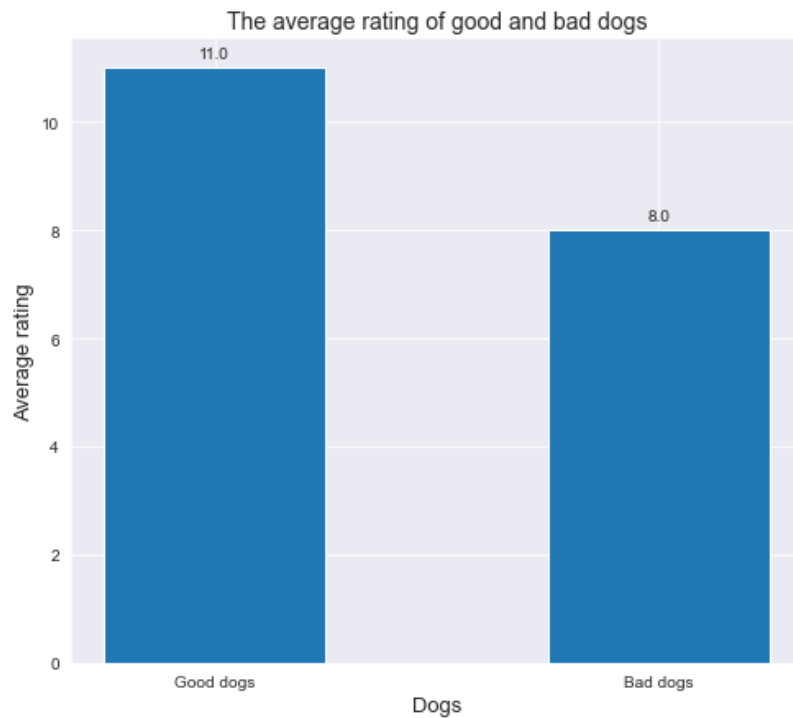
### Insight:

From the scatter plot above, there is a positive correlation between the dog ratings and total number of likes among dog lovers. The higher the dog rating, the higher the number of likes. The dog likes went to the highest levels with dogs with ratings of 10 and above (good dogs).

In fact, the likes went as high as over 120,000 likes for a few dogs with 13 over 10 ratings. This means that dog lovers tend to like the tweets about good dogs more than they did with bad dogs. This also confirms that the sentiment of the dog lovers agrees with the dog poster ratings. That the dogs are really good dogs.

### What is the average rating of both groups of dogs?

To answer this question, I will plot the average ratings of both group of dogs on a bar chart.



### Insight:

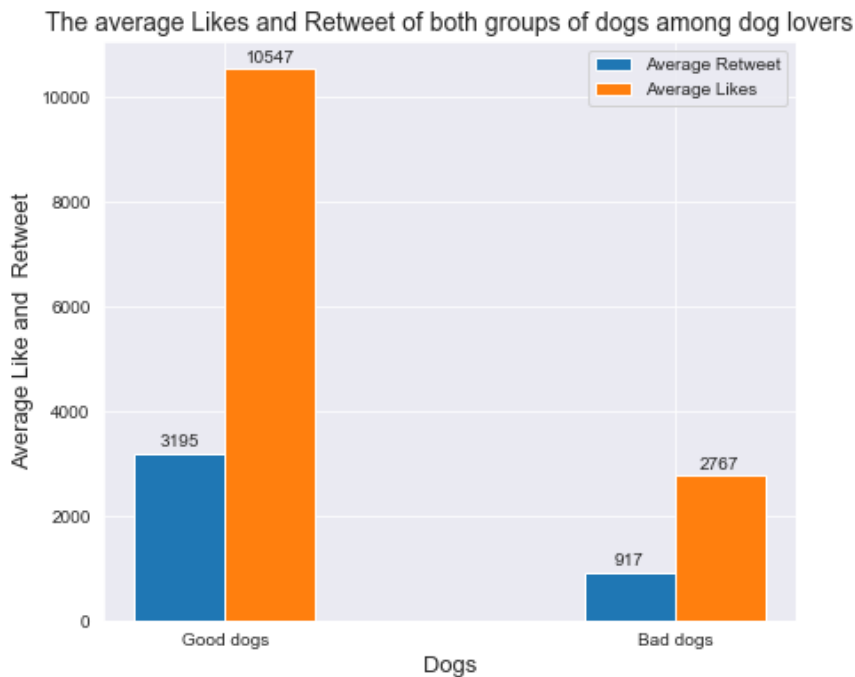
From the bar chart visualization above, it can be deduced that the good dogs have an average rating of 11 while bad dogs have an average rating of 8.

Good dogs are really good because they have ratings above 10. Bad dogs are bad because they have an average rating below 10. Again, please note that bad dogs are not really bad but this has to do with rating system.

**Which group of dogs has more likes among dog lovers?**

**Which group of dogs did dog lovers retweeted more?**

To know the group of dogs that dog lovers liked and tweeted more, I will plot the average likes and retweets of both groups on a bar chart.



### Insight:

From the bar chart, dog lovers retweeted good dogs more than bad dogs. Good dogs have an average retweet of 3,195 while only 917 retweets for bad dogs. The difference is very clear. Dog lovers seem to agree with the opinion of the dog rater again.

It can also be deduced that dog lovers liked good dogs more than bad dogs. Good dogs have an average likes of 10,547 while 2767 likes for bad dogs. The difference is very clear. Dog lovers seem to agree with the opinion of the dog rater.

Remember it's the dog poster that determines which dog is good and which is bad. He gives rating 10 and above for good dogs and rates dogs that are not too appealing to him a rating under 10.

## What breed is the dog with the highest number of likes?

### Insight:

From my exploration, I found that dog lovers liked a dog which is of the Lakeland Terrier breed the most. The dog has a rating of 13. The

Note that this record is only for dog lovers that hit the like button for the dog.

Over 130,000 people liked the dog.

I also found that dogs that are of Lakeland Terrier breed are generally good dogs. This is because their average rating is above 10. Please, note that this result is based on average. There are also bad dogs among the Lakeland Terrier breed. But there are more good dogs than bad dogs among the breed.

### Image details of the most liked dog

**Dog breed:** Lakeland Terrier

**Dog URL**

[https://twitter.com/dog\\_rates/status/822872901745569793/photo/1](https://twitter.com/dog_rates/status/822872901745569793/photo/1)

**Tweet id:**

822872901745569793

**Dog Image:**



### Which dog breed did dog lovers liked and retweeted the least?

From my exploration, I found that the dog with the tweet id of 666102155909144576 was liked and retweeted the least by dog lovers. The dog belong to the English Setter breed.

From my exploration, I found that dogs that are of English Setter breed are generally bad dogs. This is because their average rating is below 10.

Please, note that this result is based on average. It does not mean there are no good dogs among them. There are definitely good dogs among them. But it appears there are more bad dogs than good dogs among the breed. In other words, there are more dogs with ratings below 10 than dogs with ratings of 10 and above.

### **Why did the the dog has the least number of likes and retweets?**

I found that this dog has a rating of 11. That means it's a good dog. I was curious to know why the dog has such low reactions from dog lovers. I then explored more. From my findings, I found that the dog was posted the same month the WeRateDog account was created. Lower number of followers could be the reason the dog has such low number of likes and retweets.

### **Image details:**

**Tweet\_id:** 666102155909144576

**URL:** [https://twitter.com/dog\\_rates/status/666102155909144576/photo/1](https://twitter.com/dog_rates/status/666102155909144576/photo/1)

**Dog breed:** English Setter





**Conclusions:**

Dogs with high ratings tend to be liked and retweeted more by dog lovers.

Dogs with low ratings are retweeted less by dog lovers.

Dog lovers sentiment seem to agree with the ratings of the dog poster.

It appears that if the dog rating is very low or below 10, it's more likely to be retweeted and liked less by dog lovers.

**Limitation:**

In this project, I did not explore the date in which dogs ratings were tweeted.

It's important to know that the number of followers WeRateDogs have at the time of tweeting could impact the number of likes and retweets.

Dogs posted in 2016 and 2017 are more likely to have more likes than dogs posted in 2015. That's because the number of followers has multiplied over the years.

Dogs posted in 2015 are more likely to receive low reactions because the account was created in the same year. Hence, I couldn't draw conclusions on such dogs.

I must also mention that my analysis is limited and do not draw absolute conclusion on why dogs receives more reaction than other dogs. Further analysis needs to be carried out through advanced processes such as machine learning in order to draw conclusion on how dogs with low reactions differ from dogs with high reactions from dog lovers.