

# Report: Analyzing and Visualizing Data

The goal of the Wrangle and Analyze Data project is to wrangle WeRateDogs Twitter data to create interesting and trustworthy analyses and visualizations. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. It has a unique rating system, which is a big part of the popularity of WeRateDogs.

In the analyzing and visualizing data section of the project, three separate insights are generated with visualizations. Also, the data used to produce the visualizations are clearly documented. This report communicates all the insights and displays the visualizations produced in the Wrangle and Analyze Data project.

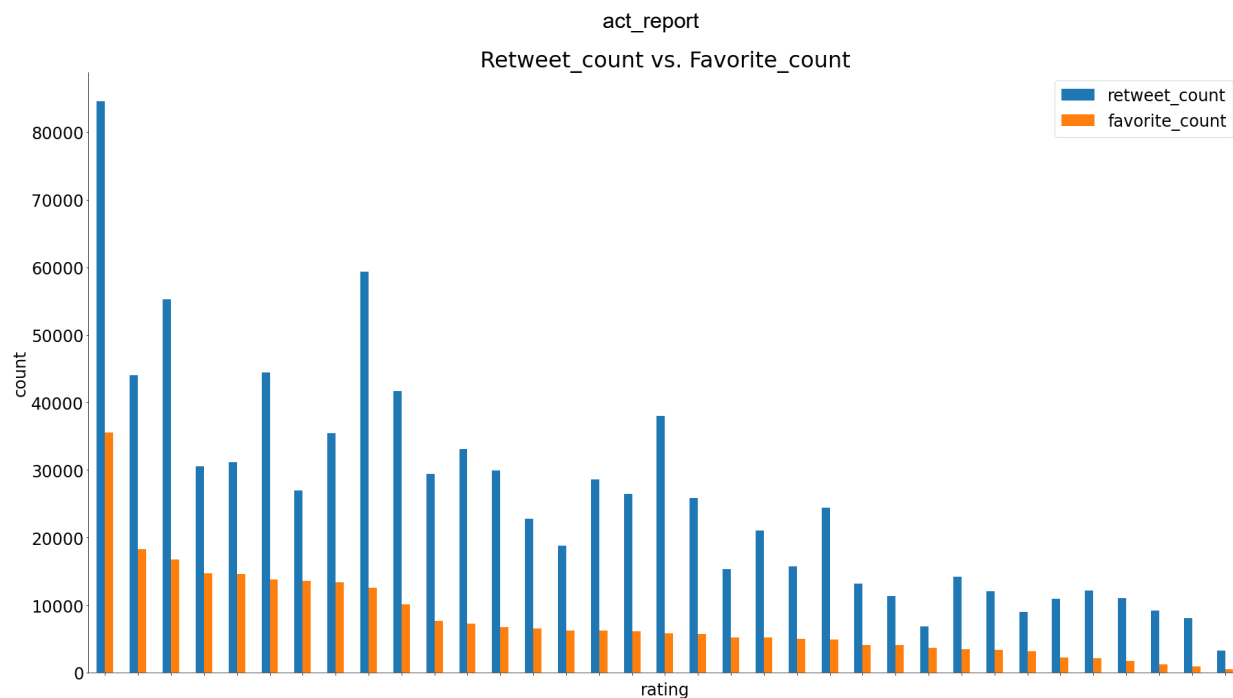
## Insight 1: Which dog is the most popular?

In the WeRateDog Twitter account, each tweet receives numerous likes which is termed as favorite counts. Moreover, each tweet becomes retweeted many many times. Based on the favorite count and retweet count data, the most popular dog is identified. The dog that has obtained the maximum favorite count (72901) is a Labrador Retriever. The tweet became retweeted for 148415 times which is the highest number of retweets for any dog. According to the twitter data, this doggo has received one of the top most ratings (1.3) too. The name of the doggo is missing in the dataset; however, it has a lovely photo which can be seen below.



## **Insight 2: What is the correlation between favorite count and retweet count?**

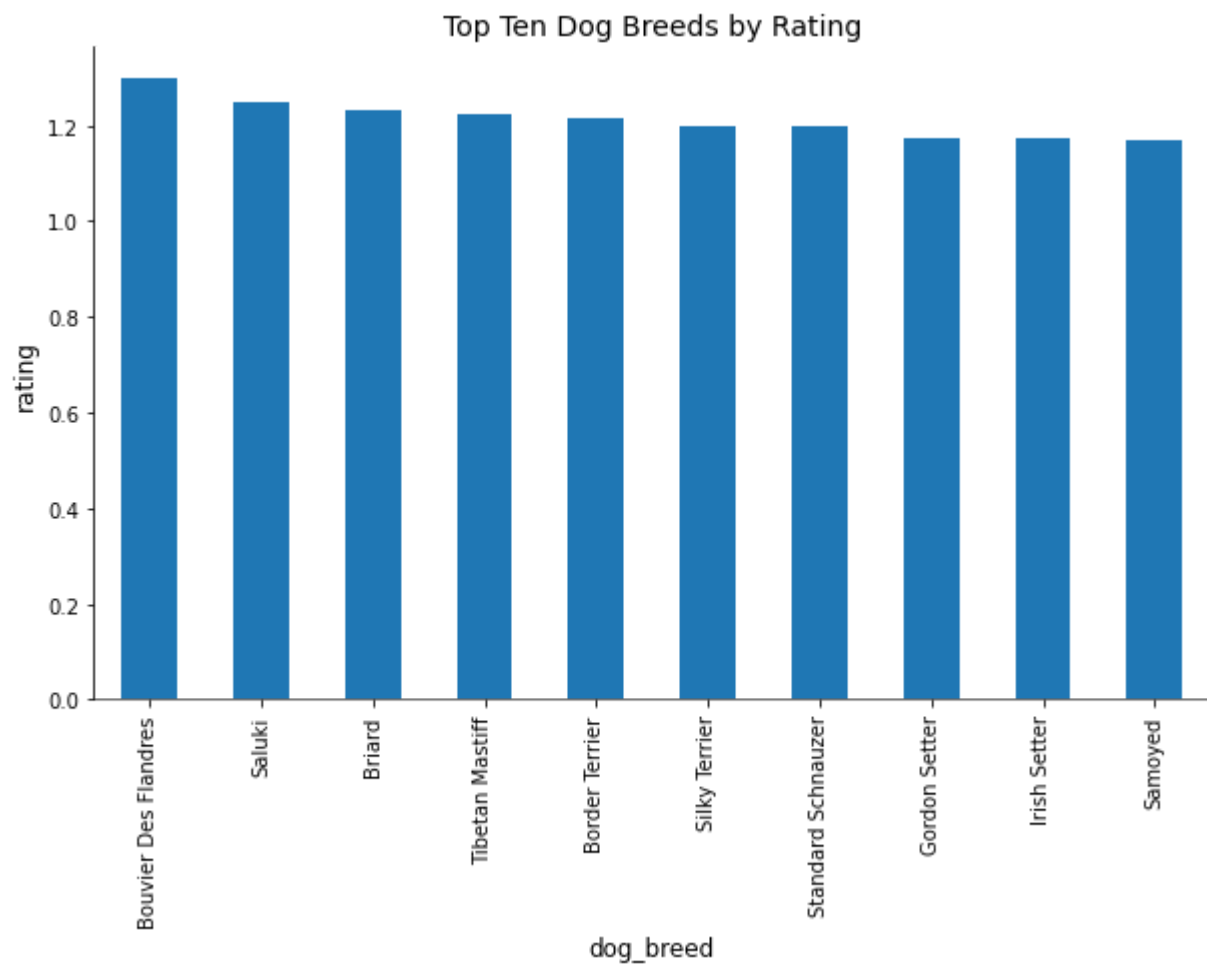
To answer this question, first, we have identified the top most rating (1.4) that present in the dataset. In the next step, all the tweets are filtered for this 1.4 rating. With that filtered data, we investigated the relationship between the number of favorite count and retweet count.



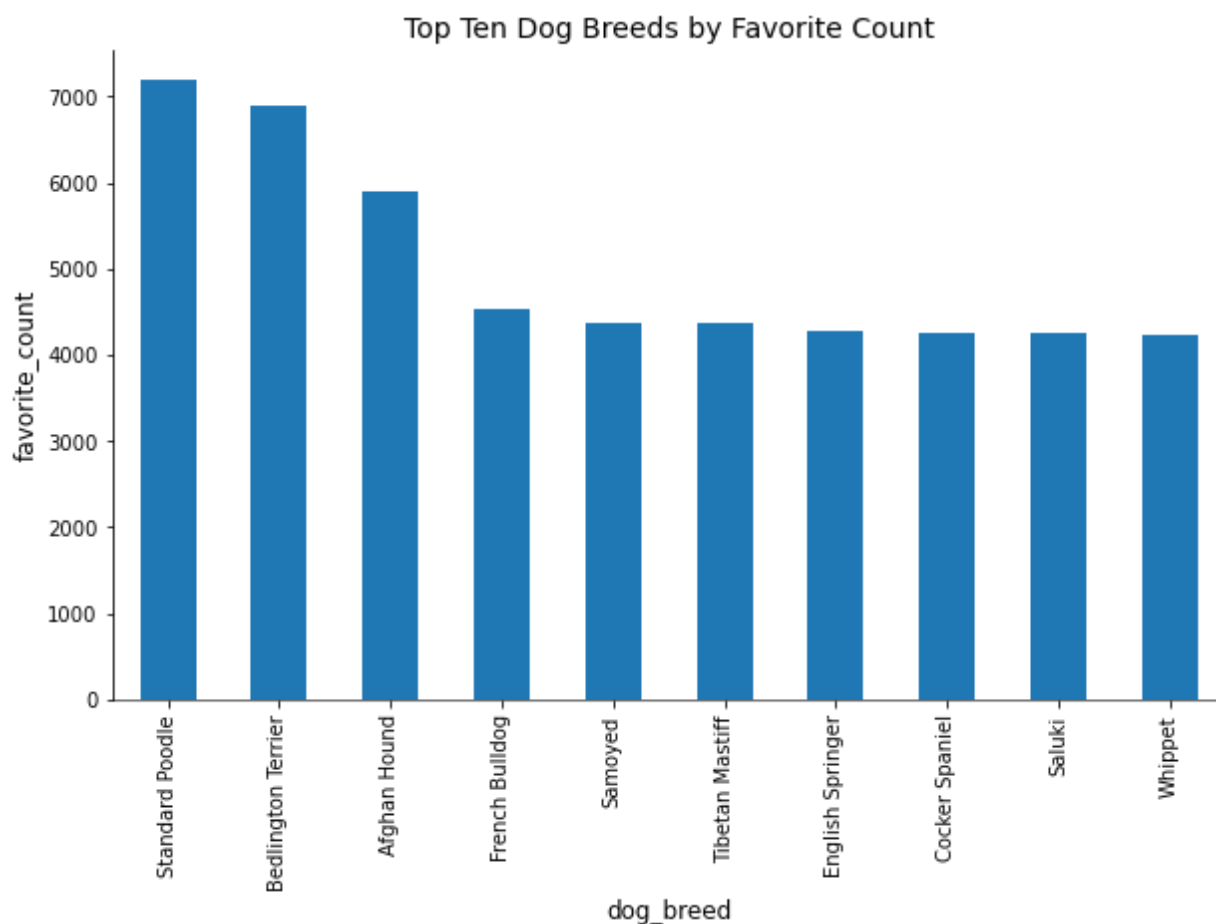
For a specific rating, favorite count and retweet count values are found proportionate. That means, a higher number of retweets will usually generate a higher amount of likes.

### Insight 3: What are the top ten breeds of dogs?

The top ten breeds of dogs are identified by rating and then, by favorite count. Mean values of rating and favorite count are used to generate this insight. The results are displayed in the following bar graphs.



Bouvier Des Flandres dogs have received the maximum rating(1.3), followed by Saluki (1.25) and Briard (1.23) in the second and third places, respectively.



Standard Poodle breed dogs have received the most likes (7184). Bedlington Terrier and Afgan Hound breed are following with 6886 and 5895 counts, respectively.

Most interestingly, the top ten breeds determined by rating do not actually received so many likes, if we compare the results with the top ten breeds by favorite count. For instance, Bouvier Des Flandres, the highest rated breed is not seen in the top ten list by favorite count.

In [ ]: