

# WeRateDogs Archive

## Data Analysis

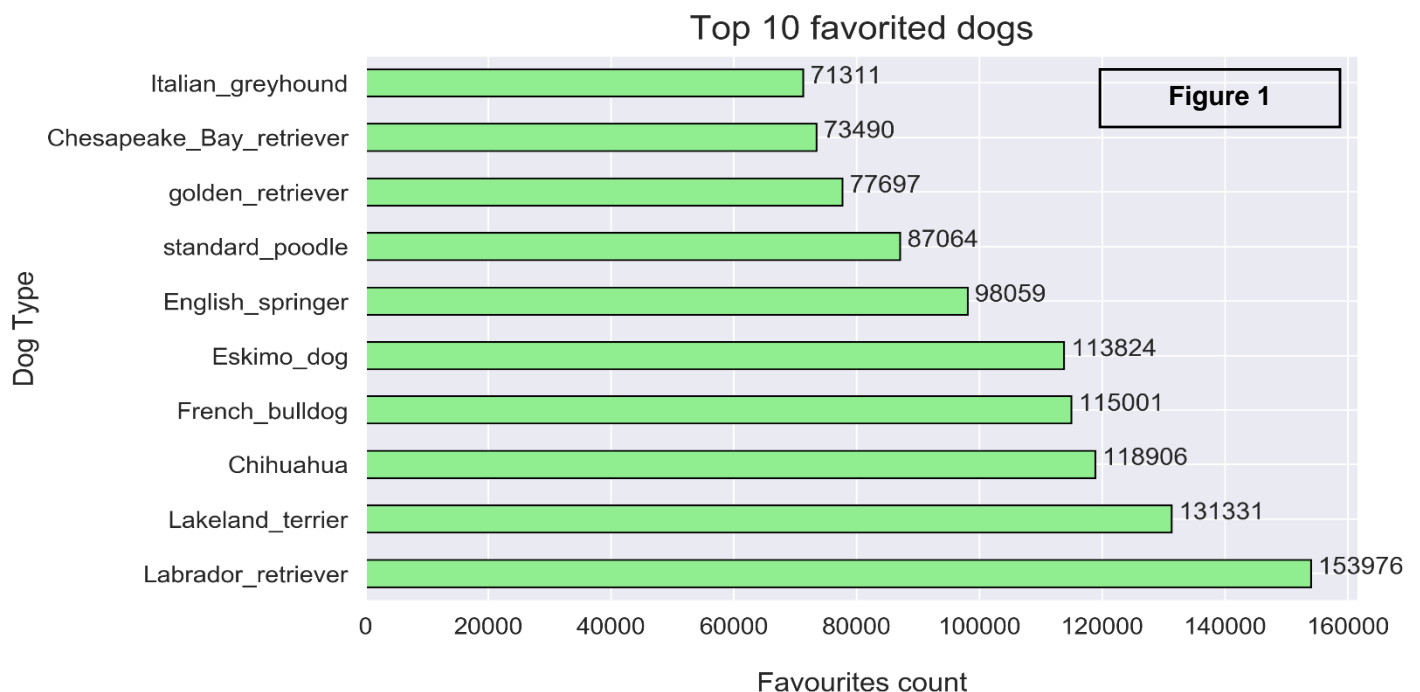
### 1) Which dog type gets the highest favorite count?

The WeRateDogs archive contains so many dog breeds (about 1750 breed). According to the data, it was possible to know exactly the favorite count of each dog breed and then sort them in descending order (shown in figure 1). For more clarity, I only showed the “Top 10 favorited dogs” as follows:

- Labrador retriever (The king of list)
- Lakeland terrier
- Chihuahua
- French bulldog
- Eskimo dog
- English springer
- Standard poodle
- Golden retriever
- Chesapeake Bay retriever
- Italian greyhound



Labrador Retriever

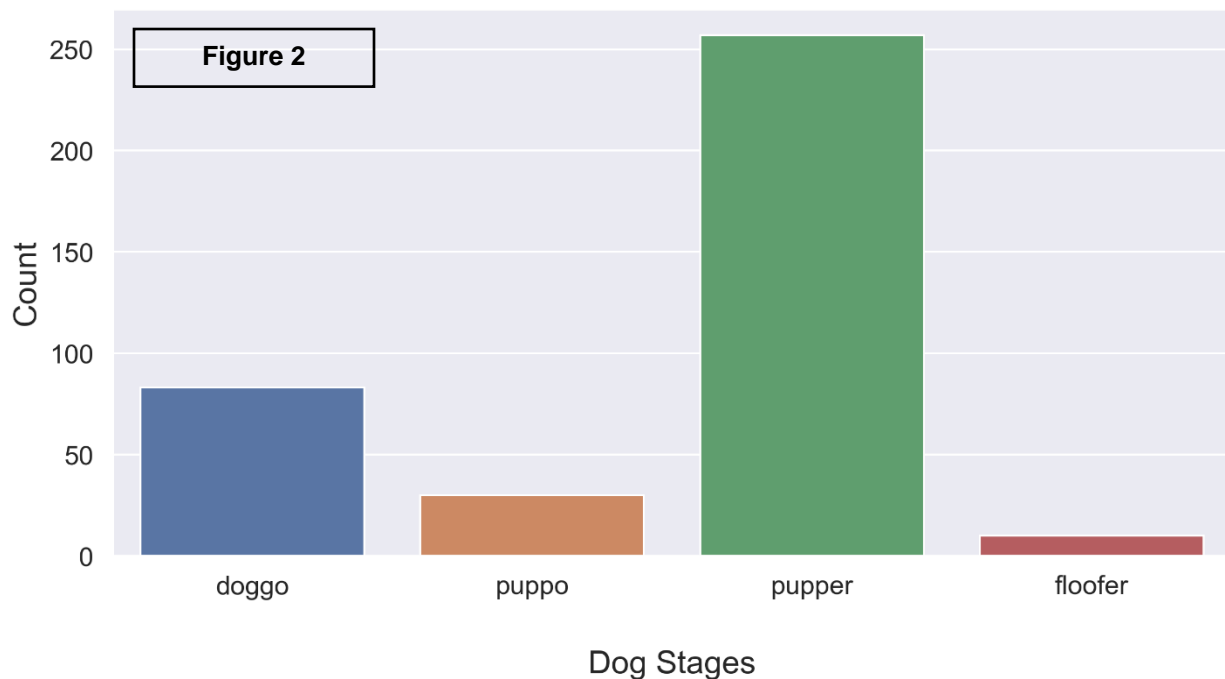


## 2) which dog stage gets the highest retweet count?

In WeRateDogs, there are four main dog stages as follows:

- **Doggo**: this can technically be a dog of any size, but it is most often used to denote a dog of medium to large size.
- **Pupper**: in contrast to "doggo," "pupper" usually means a dog of smaller size.
- **Puppo**: An alternative word to use when seeing a cute dog or puppy.
- **Floofer**: Woofers are large doggos and a floofer is a big very hairy Woofer

Although most of dogs on the archive are not classified into one of the four stages above (about 1973), there are 257 puppers, 83 doggos, 30 puppos and only 10 floofers (figure 2).

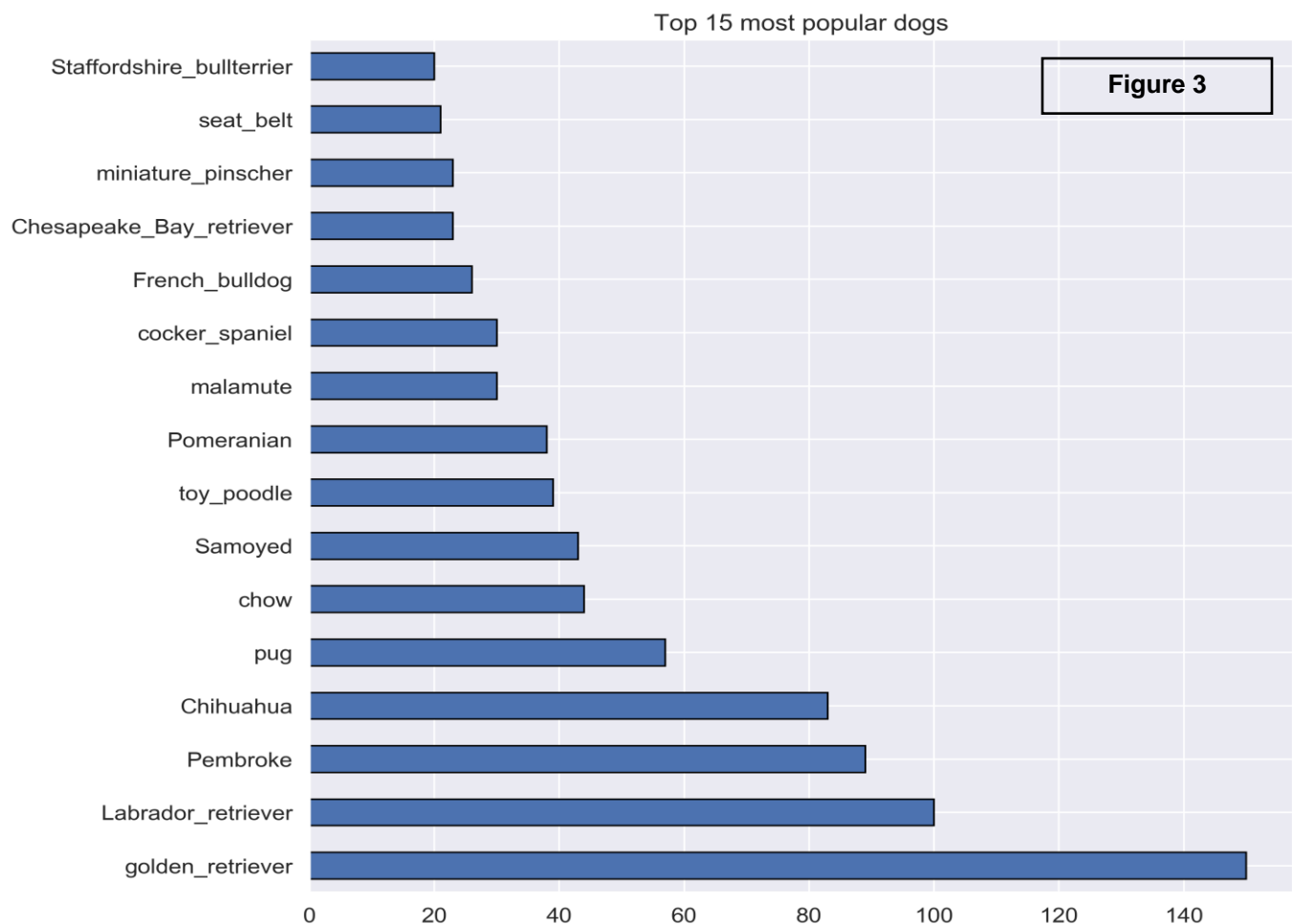


### 3) What are the most and the least popular dogs on WeRateDogs?

From the gathered data, a neural network was able to identify each dog breed for each tweet in WeRateDogs archive. By simple data analysis, it was easy to find the most common dog breed on WeRateDogs (also the least) as follows:

- **The most common:** Golden retriever
- Labrador retriever
- Pembroke
- Chihuahua
- Pug
- .
- .
- .
- **The least common:** there are many dogs that were detected only once like: Cougar, timber-wolf, canoe

In the figure 3 below, more details about the top 15 most common breeds are shown.



#### 4) What is the accuracy of the neural network predictions?

As previously mentioned, a neural network was used to detect the dog breed in the image of each tweet. However, as it is expected from any AI system (till now), there is always an error percentage. By analyzing the correctness of the predicted data, it was found that approximately 87.54% of the predictions resulted in real dog breeds and only 12.46% predicted wrong things (not dog breeds) as shown in figure 4.

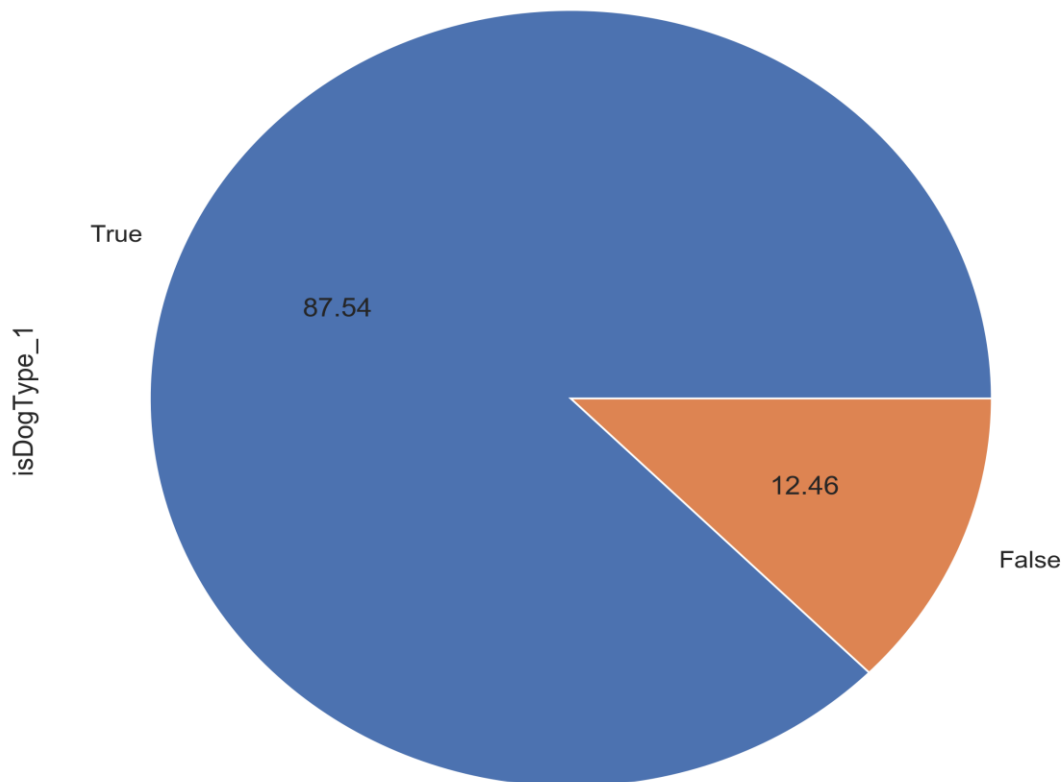


Figure 4