

WeRateDogs Project

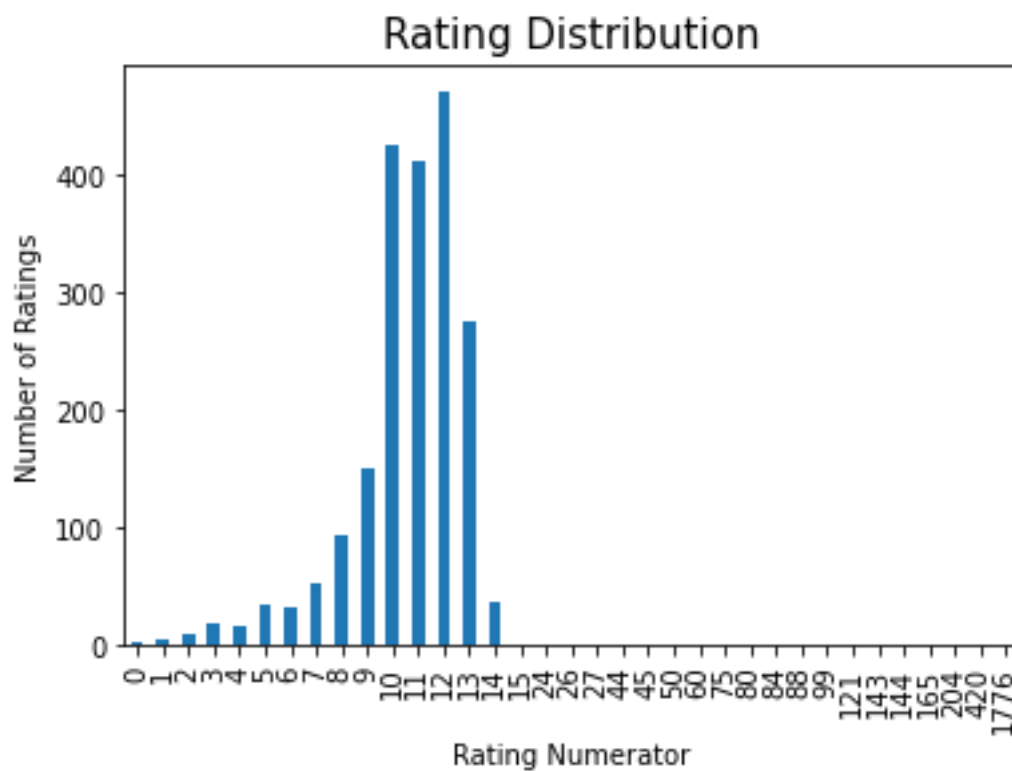
Analyzing and Visualizing Data

Table of contents:

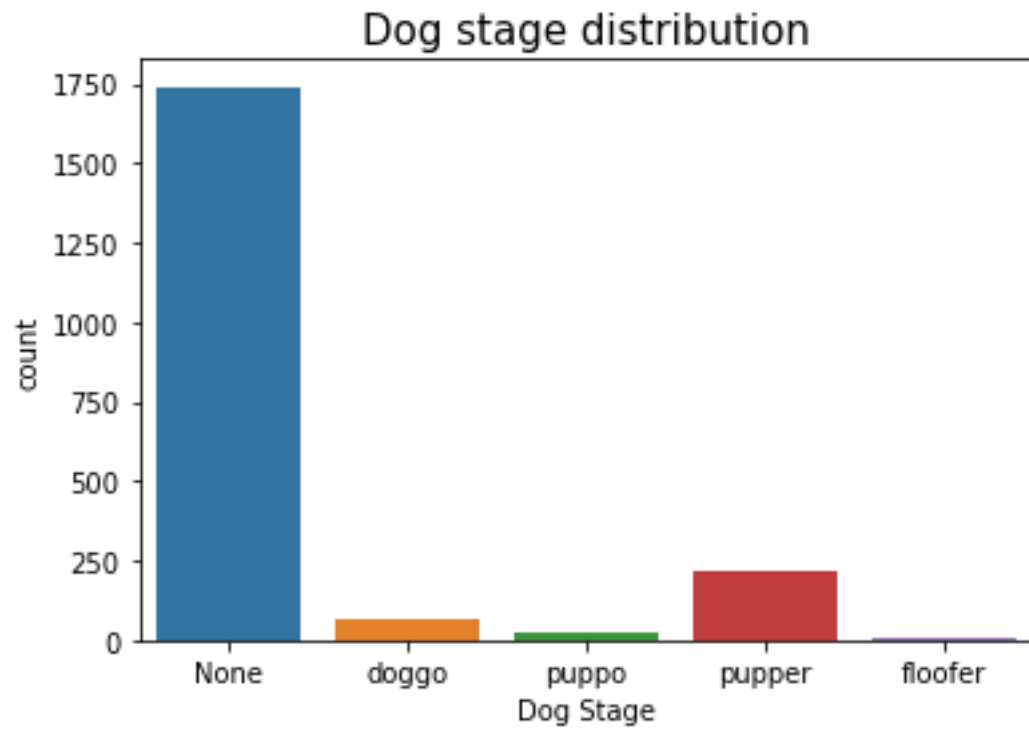
- Rating distribution
- Dog stage distribution
- Breed and prediction distribution
- Which dog stage has the highest retweets and favorite
- Source distribution
- Relation distribution between retweets and favorites

Before starting with my analysis I checked the 'twitter_archive_master.csv' file using .info() and changed the datatype for (source, tweet_id, timestamp). Then used .describe().

Rating distribution

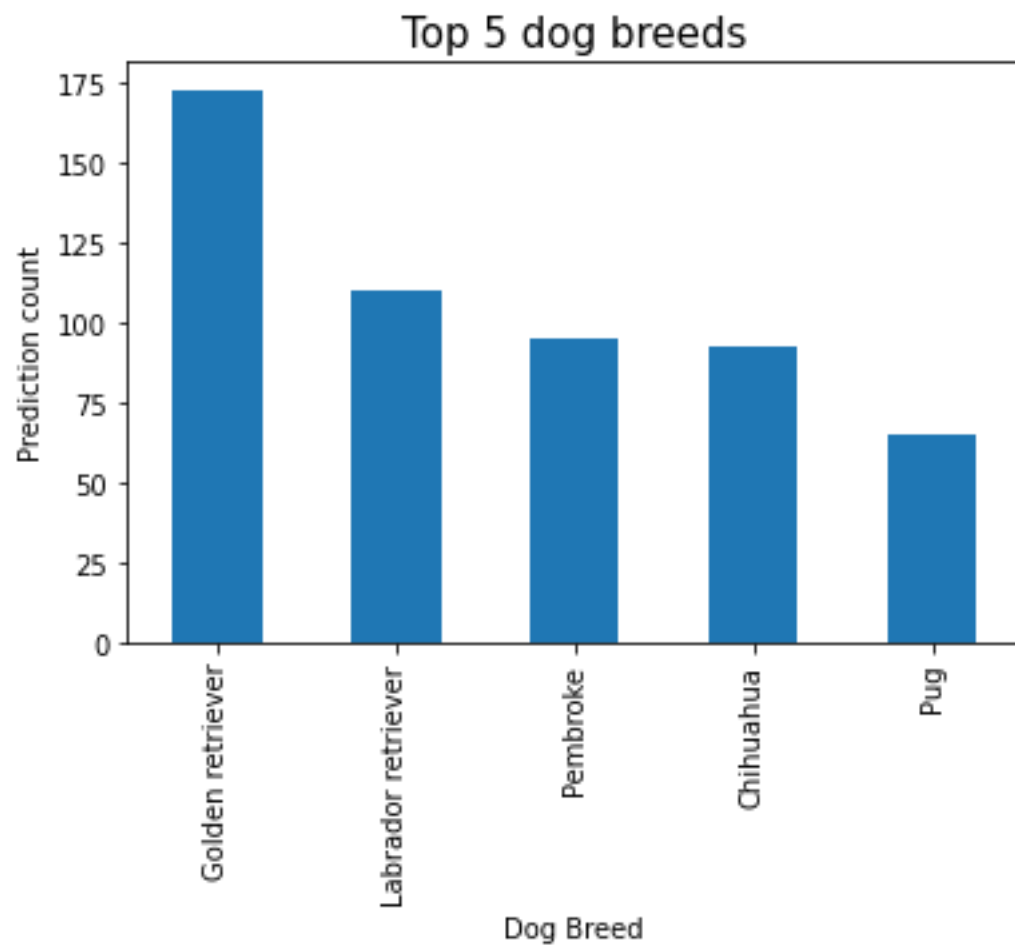


Dog stage distribution

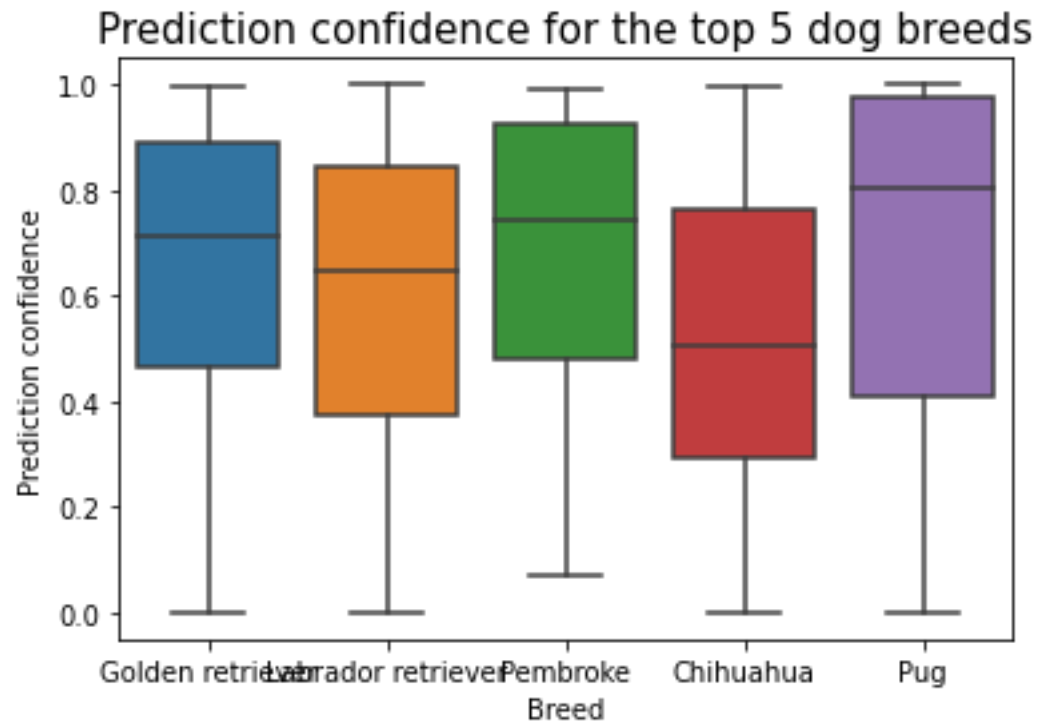


As the bar chart shows, the most common owned dog stage is pupper, the None stage might be dogs for sale. This shows that most people prefer to adopt a dog at a young age.

Breed and prediction distribution

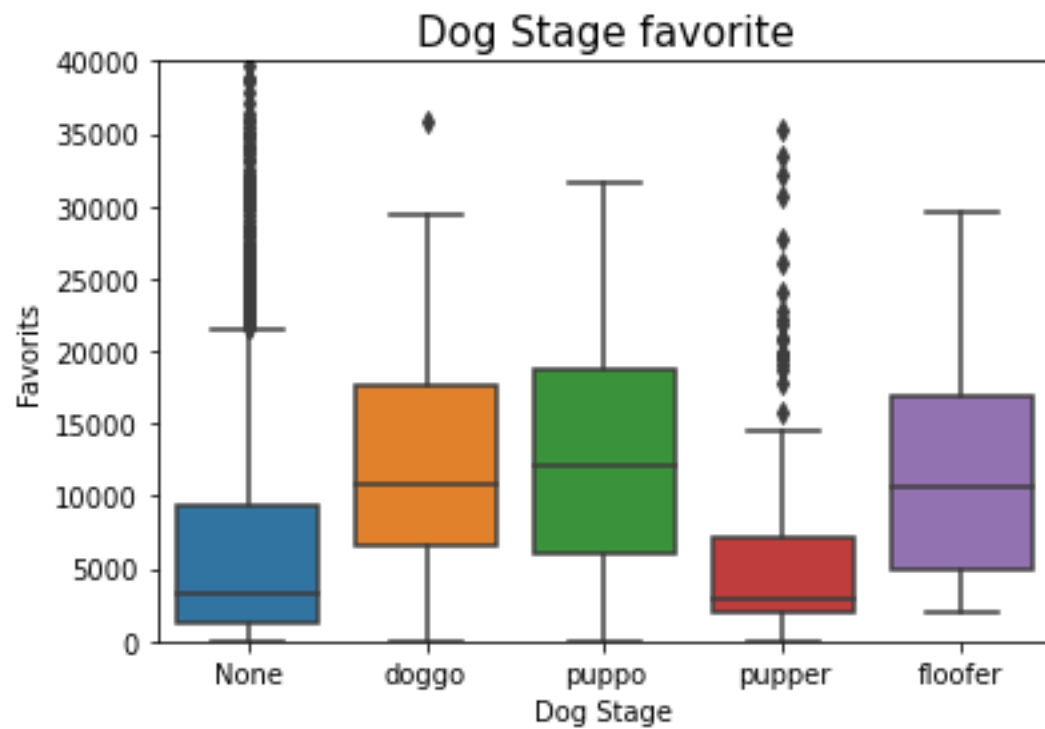
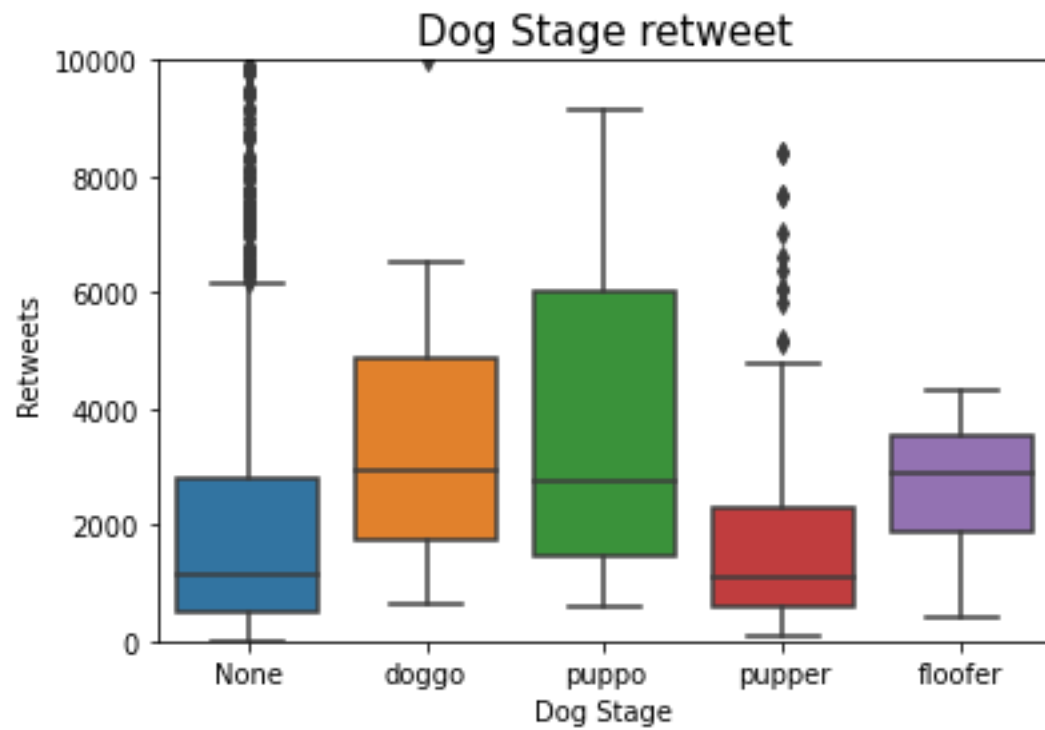


As the bar chart shows, the highest dog breed prediction is Golden retriever. Maybe this also implies that the most common dog breed is Golden retriever. Golden retrievers are known for their friendly attitude especially towards children.



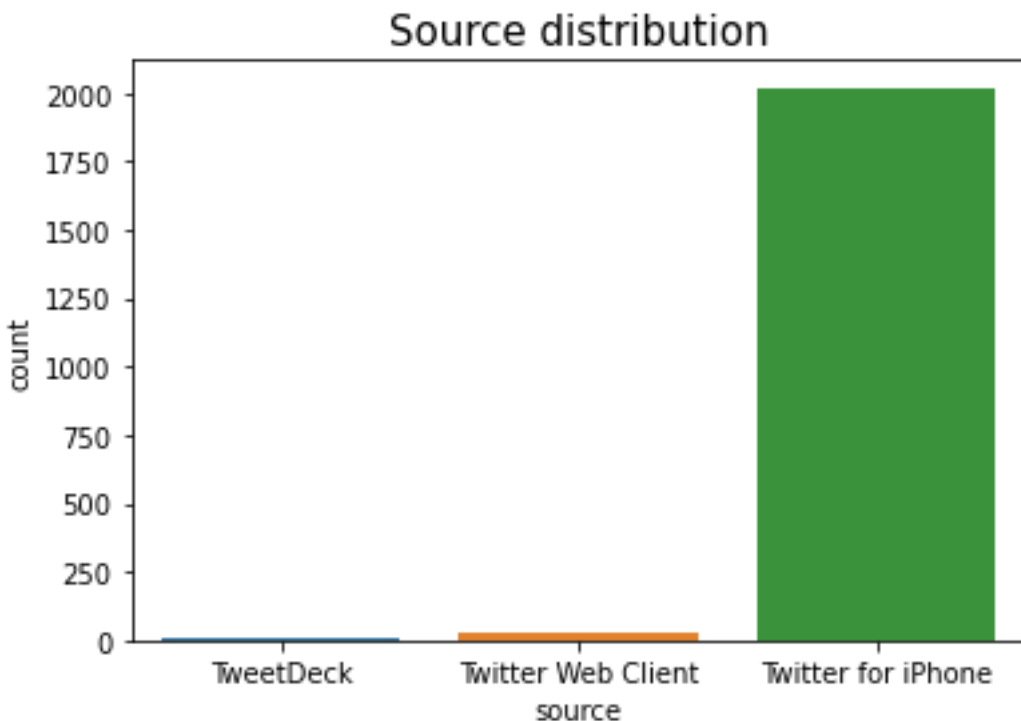
As the boxplot shows the highest confidence rate was for Pug and Pembroke, while the rest breeds were unsure like Chihuahua. Maybe this is because Pugs and Pembrokes has more distinguished features than other dogs.

Which dog stage has the highest retweets and favorite



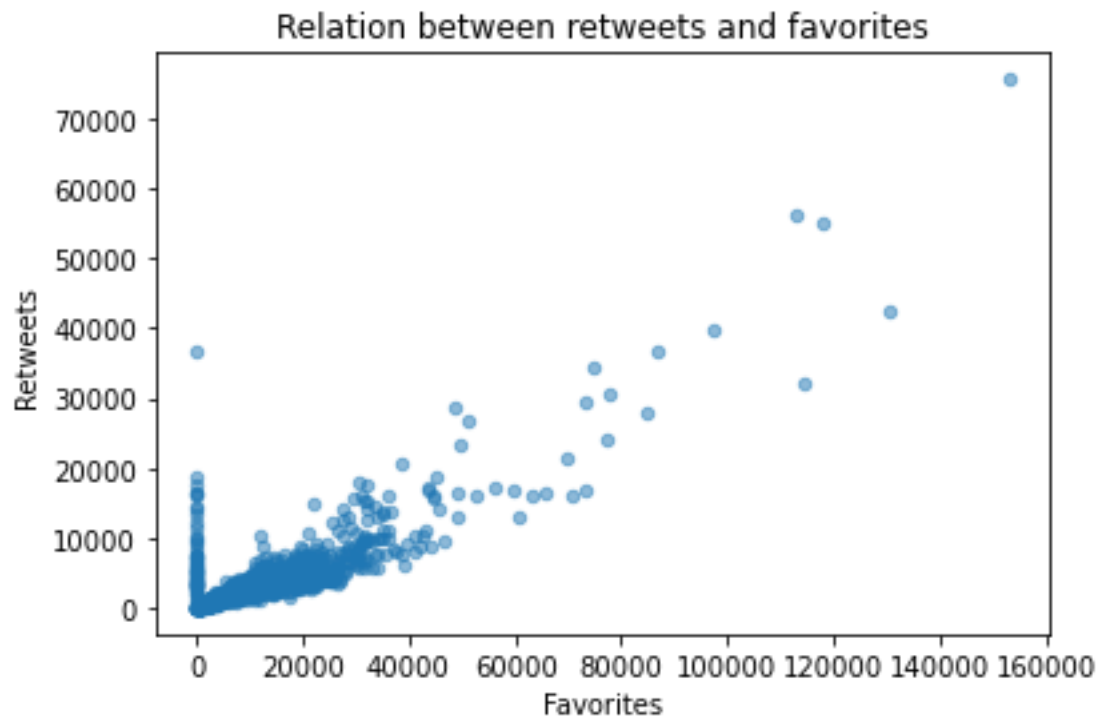
As the two boxplots shows the Puppo stage has the most retweet and favorite counts. This also confirms that it is preferable to adopt a dog at a young age. On a side note the None stage which might imply to dogs for sale or adoption has the hightes retweets and favorites which shows that WeRateDogs can be a recomended place be users to adopt or buy a dog from previous owners. WeRateDogs not only for dog owners to showoff their dogs and rate others but also for adopting new dogs.

Source distribution



As the bar chart shows, the most common source was the Twitter for iPhone. that means that users use Twitter app on iPhone to upload pictures tweet, retweet, and favorite on WeRateDogs.

Relation distribution between retweets and favorites



As the scatter plot shows, there is a positive correlation between retweets and favourites. That means that the highest a dog gets retweeted the higher it gets favorited.