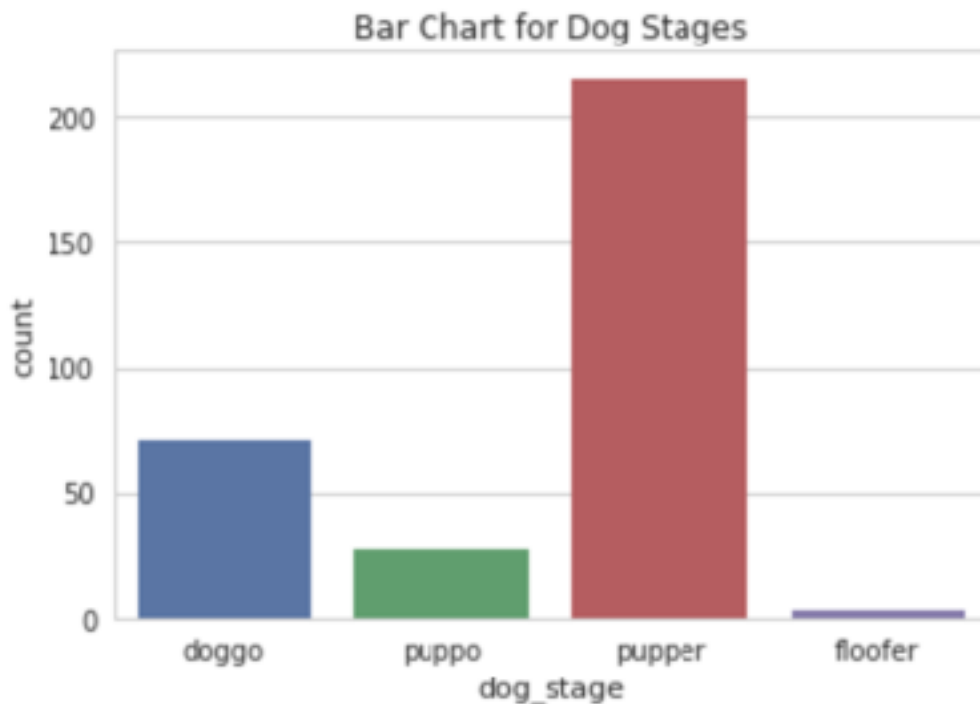


Report of 'we rate dogs' Analysis

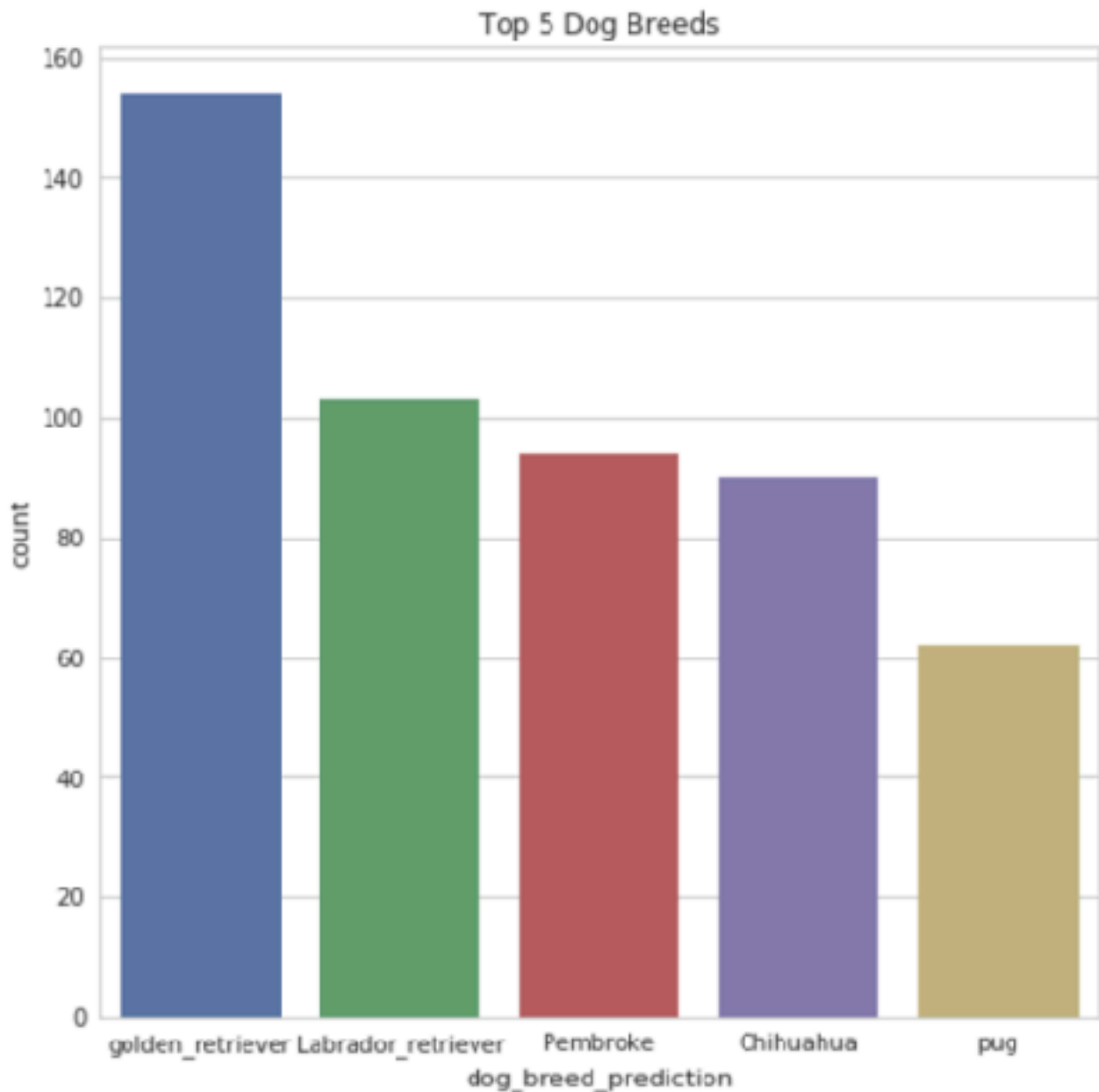
Based in our wrangling work that we have done (assessing and cleaning 'we rate dogs' data), let's analyze the dataset and make some insights and visualizations to explore the data.

1) Exploring Dog Stages:



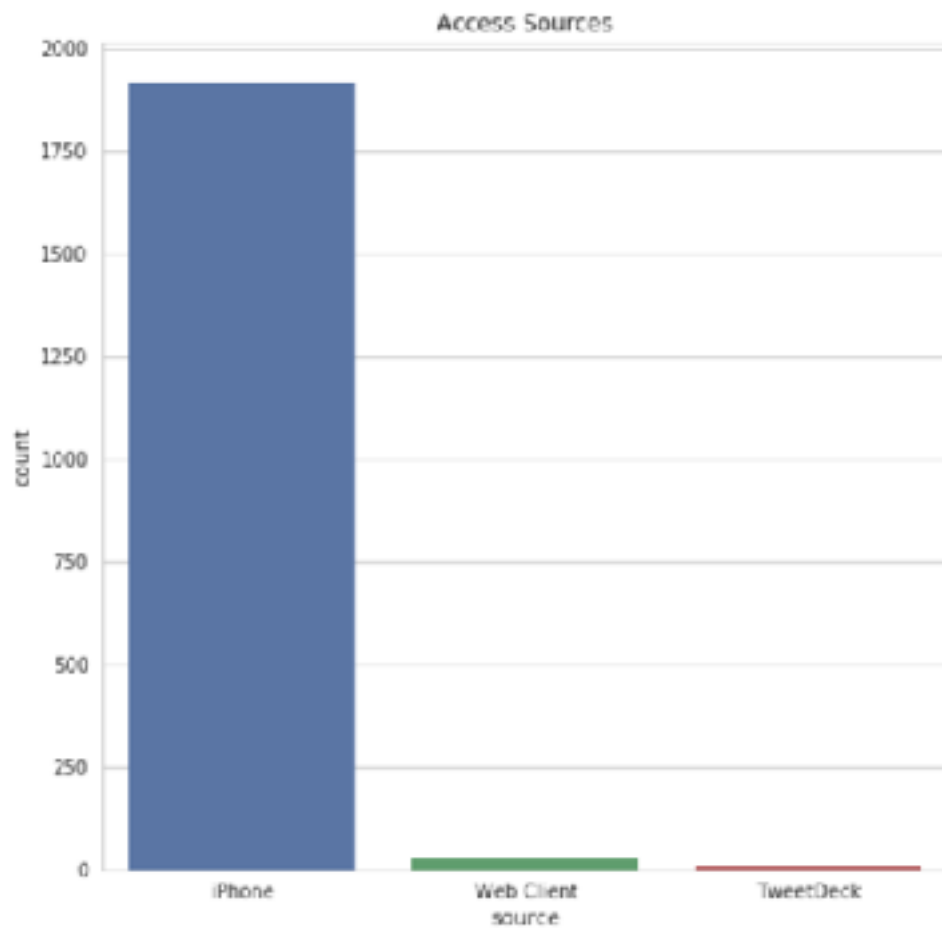
As we can see above, 'Pupper' is the most frequent stage of dog with a count of >200. While other stages are not mentioned that much in tweets. 'Doggo' stage of dog comes next, while flooder is the least mentioned stage with a count that is less than 10.

2) Exploring dog breeds



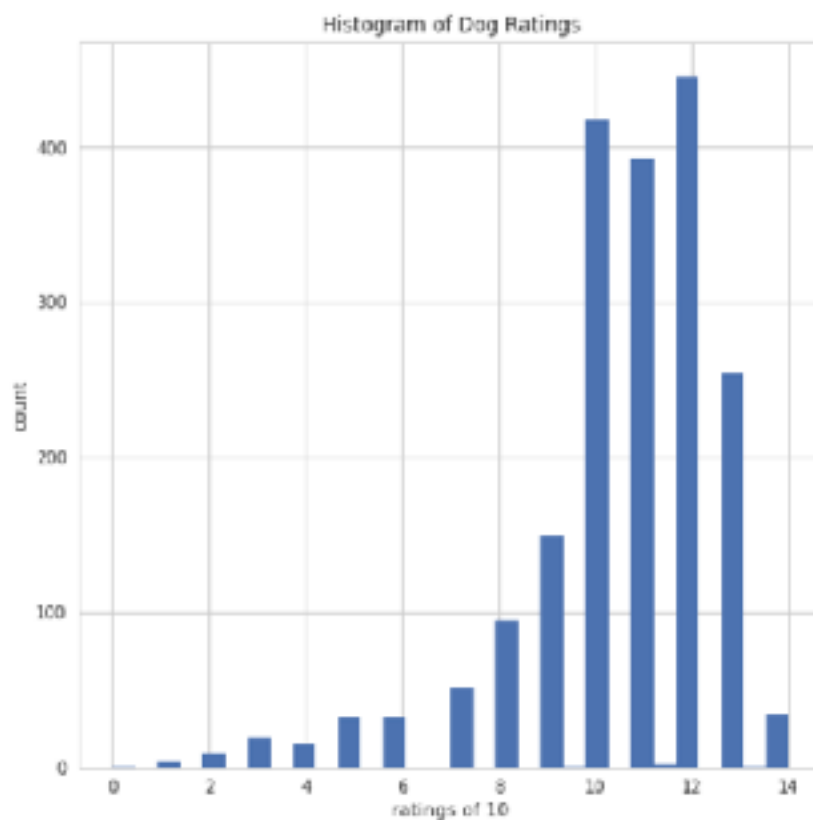
As we can see above, 'Golden Retriver' is the most frequent dog breed predicted by Neural Network with a count of more than 150, followed by 'Labrador Retriver', 'Pembroke', 'Chihuahua' and 'pug' breeds.

3) Exploring tweet sources :



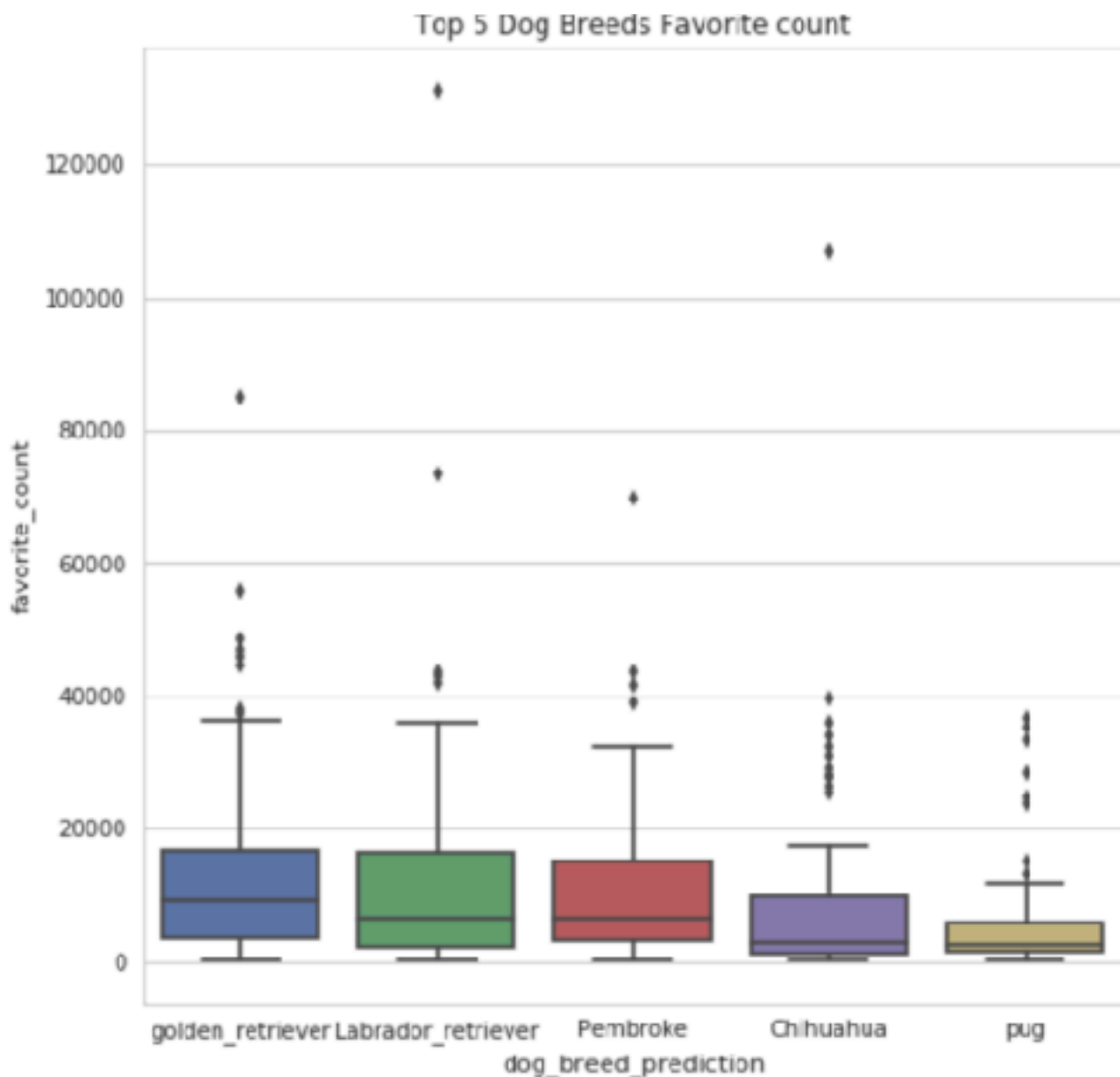
Counting the proportions of each source, I got that 98% of tweets were written from an iPhone :) while 2% of tweets came from Web client or TweetDeck.

4) Exploring Dog ratings :



More than 400 dogs were rated 12/10. Also by looking at 25% and 75% quartiles, most of dogs have been given ratings between 10 and 12. However, no ratings of more than 14 existed.

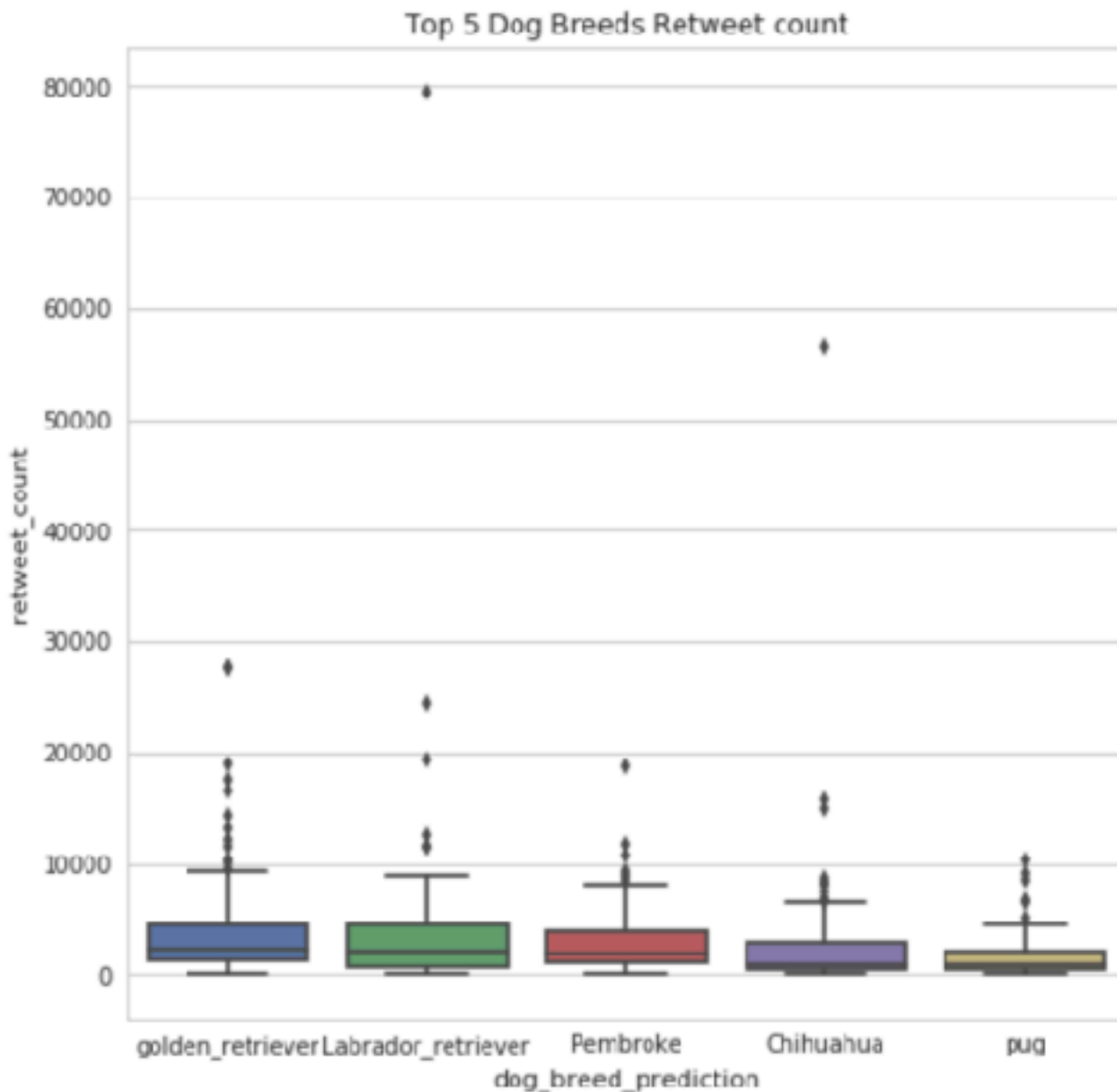
5) Exploring favorite count by dog breed :



We see from the box plot above, that favorite count for tweets mostly is below 20000 favorites. However we see some outliers. For example one tweet with “Labrador_retriever” as breed prediction for dog image got more than 120000 favorite hits! Also, we see multiple “golden retriever”

dogs got more than 40000 favorites. In conclusion, we can say that “Labrador_retriever” got higher favorite count in general.

6) Exploring retweet count by dog breed :



The box plot above shows retweet counts for most predicted breeds of dog image. Box plots are good for detecting outliers. We see in the plot that one tweet with “Labrador_retriever” dog got almost 80000 retweet hits! This is most likely the same “Labrador_retriever” tweet that got more than 120000 favorite hits in the previous plot. We can see that different breeds have similar retweet count and there’s no noticeable difference between them.

We notice the difference between this plot and the previous one. The retweet counts are generally much less than favorite counts.