WeRateDogs Tweets Archive

By Muhammed Jimoh

18th February 2023

A Samoyed

The method used to analyze the WeRateDogs Twitter archive is summarized in this report.

Data Analysis

Cleaning and merging the various datasets collected from various sources allowed for analysis to be done.

1. Reading the clean data
2. Performing quick analysis i.e. data types, count
3. Describe both numeric and categorical data types
4. Breed value counts
5. Tweet sources
6. Count ratings
7. Top 5 common dog breeds

8. Number of favorite_count for dog stages
9. Percentage of dog stages

Three main sources of the various devices used to tweet about the dogs were noticed, with iPhone devices ranking highest among them. The high resolution photos are necessary for this purpose.
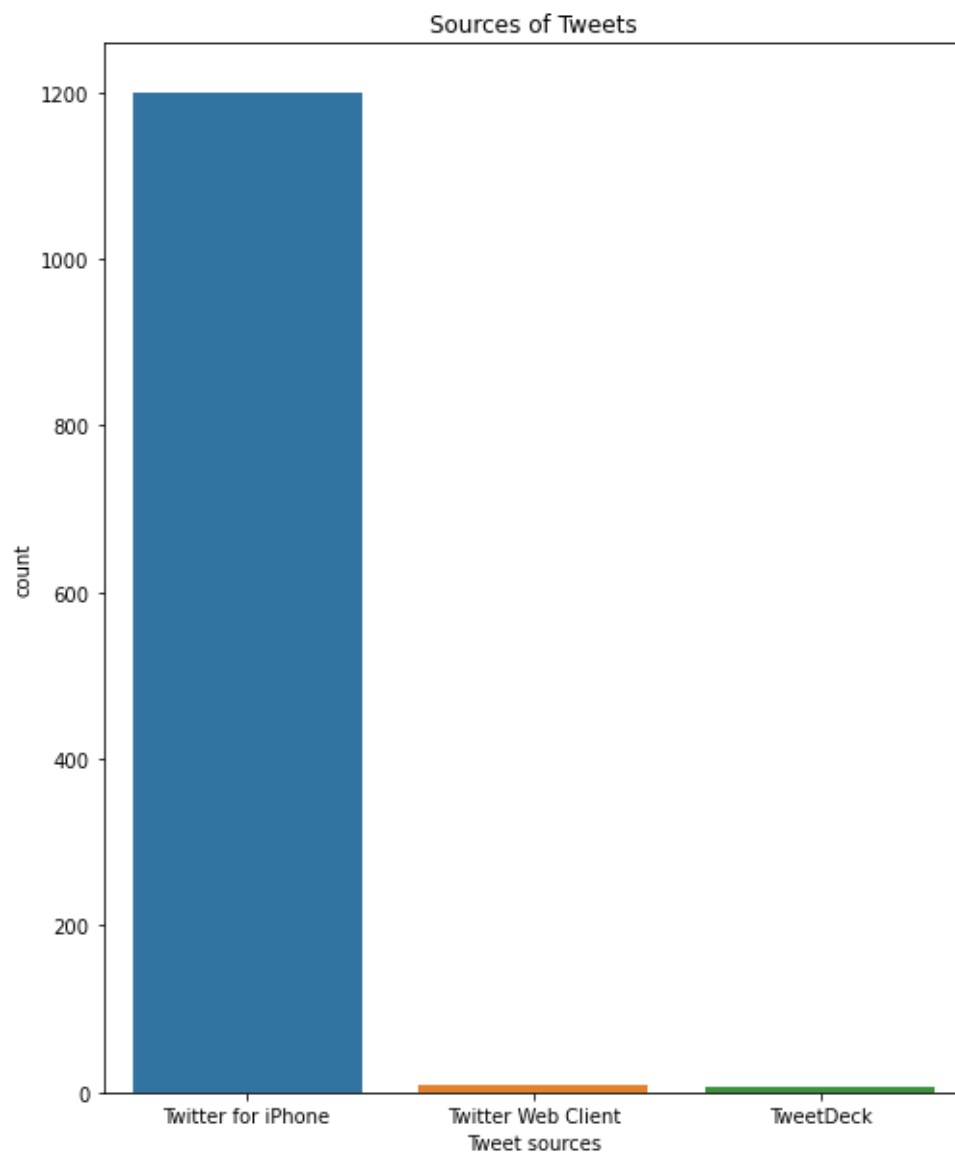

Fig 1.0

After standardization, the numerator number has a range of 0 to 15, with the majority of ratings falling between 11 and 13. Also crucial to notice is the left-skewed nature of the data.
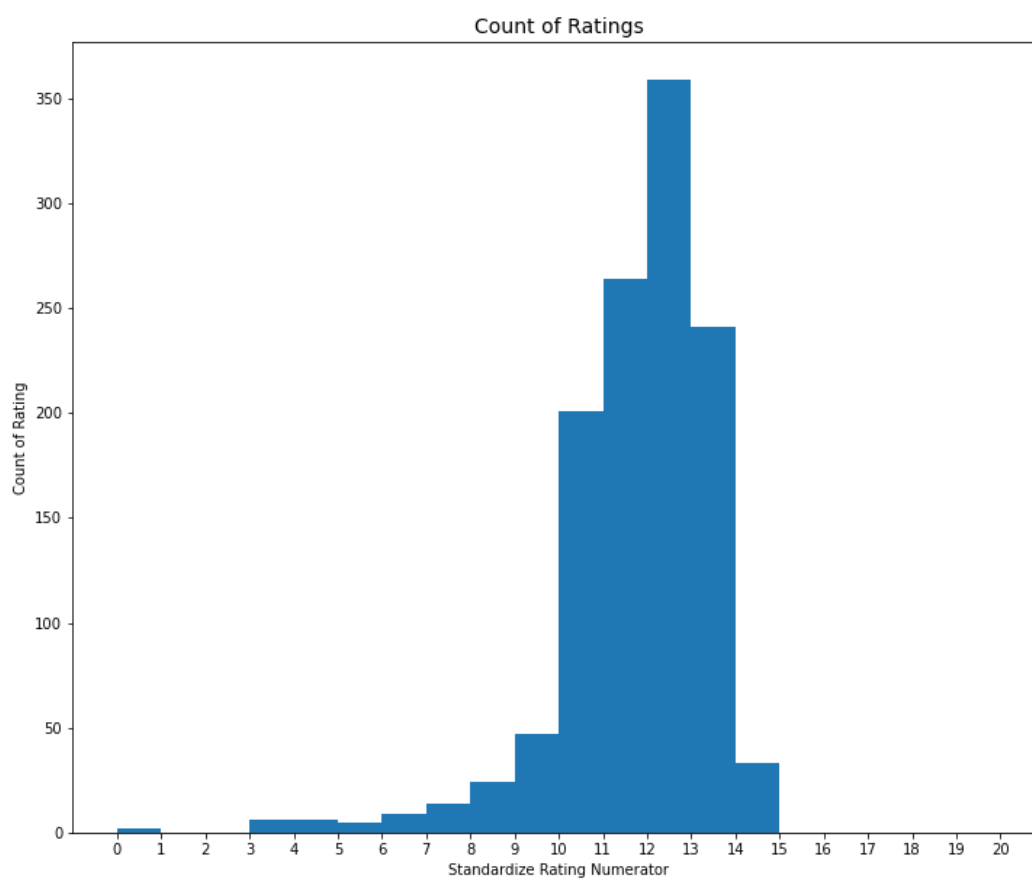

Fig 1.1

I discovered that the Golden Retriever, Pembroke, Labrador Retriever, Chihuahua, and Samoyed are the top 5 dogs that receive the most tweets.
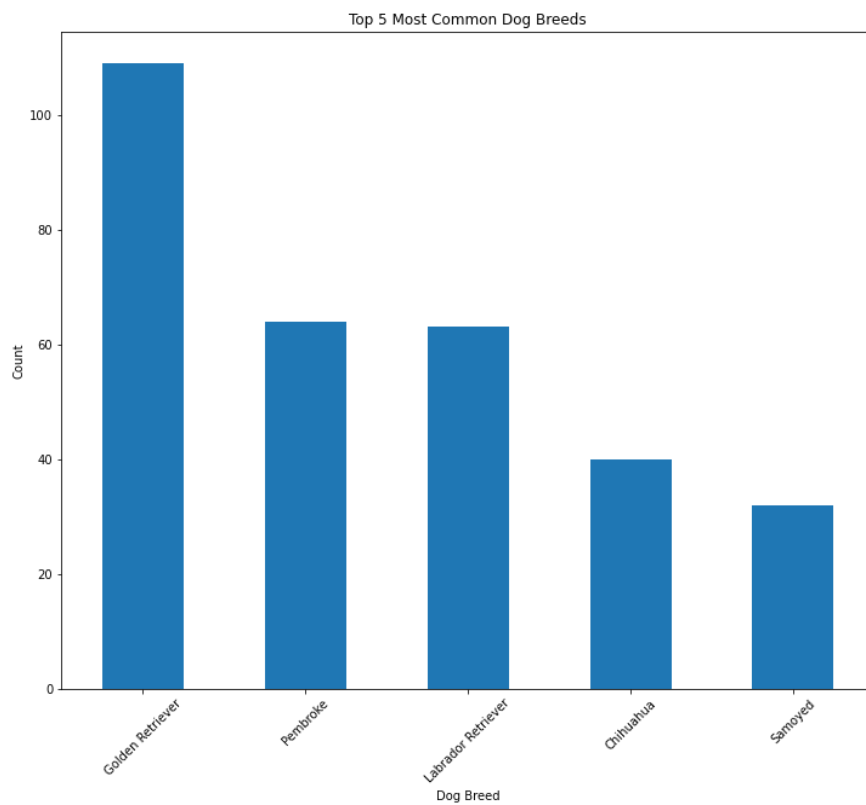

Fig 1.2

The total favorite_count for each breed of dog was also plotted, and it was discovered that doggopuppo, whose breed name is *Flat-Coated Retriever*, had the greatest favorite_count with a value of 47844.
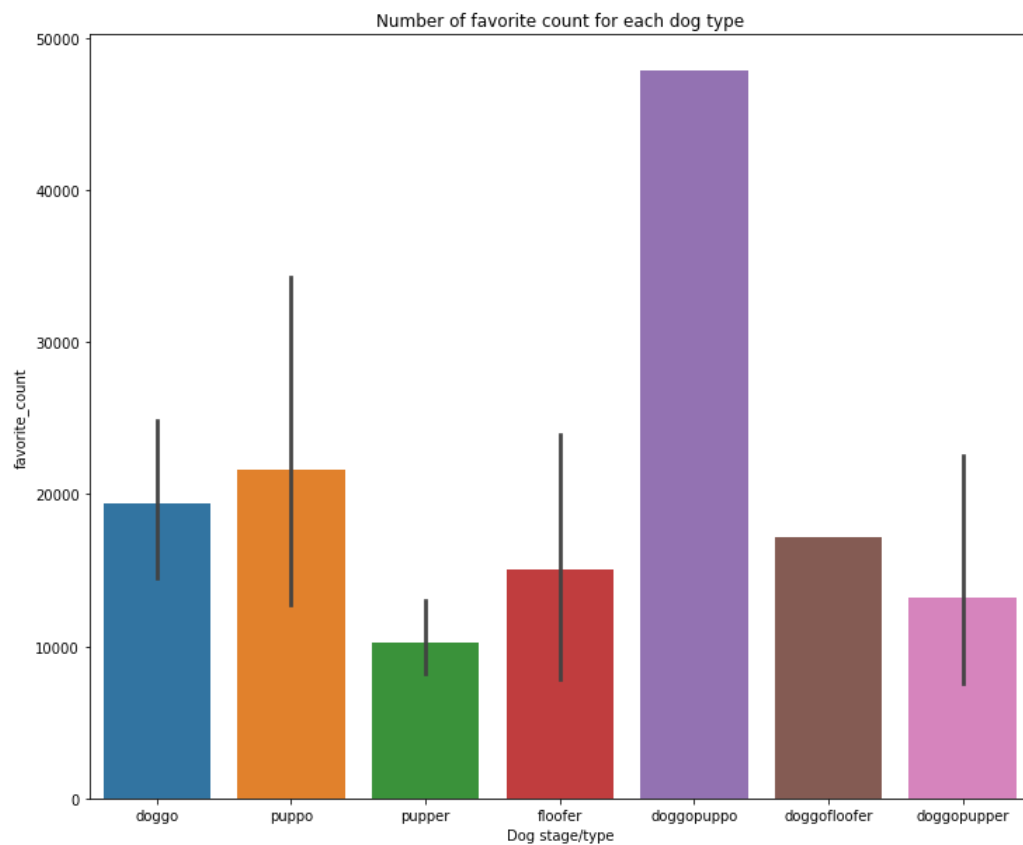

Fig 1.3

The percentage of each breed of dog was also recorded; based on my analysis, it can be seen that the pupper had the highest percentage (51.9%) and the doggofloofer and doggopuppo had the lowest percentages (0.48% each).
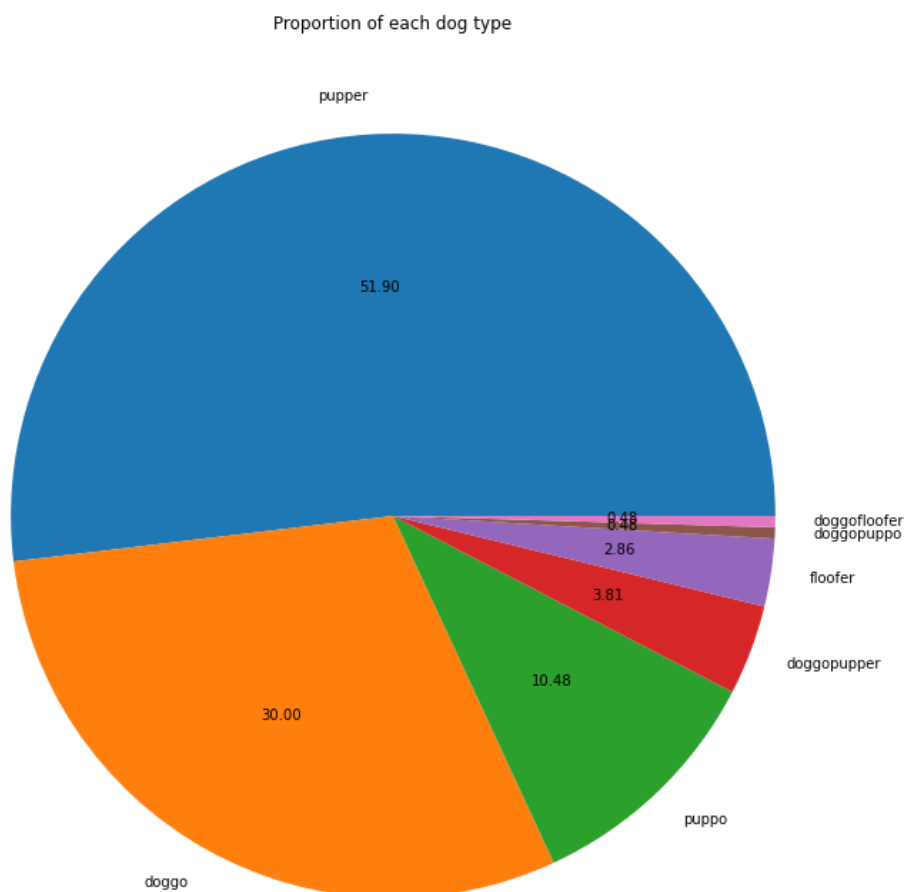

Fig 1.4