Report

Analysis of WeRateDogs Tweets Data

WeRateDogs is a Twitter account where people rate dogs out of 10. Interestingly, some of the ratings are greater than 10 which is the fun part of the accounts. The dogs rated are grouped into different stages.

Analysis was performed using the data after it was cleaned. The following insights were made from the analysis.

➢ We analyse the various dog stages (pupper, doggo, puppo, and floofer) to determine the dog stages that are likely to be rated high. Even though less than 50 puppo were rated, it has the highest ratings. Averagely, each puppo was rated at 12.00 more than the doggo and the floofer with a mean rating of 11.85 and 11.89 respectively. The pupper was least rated at an average of 10.73
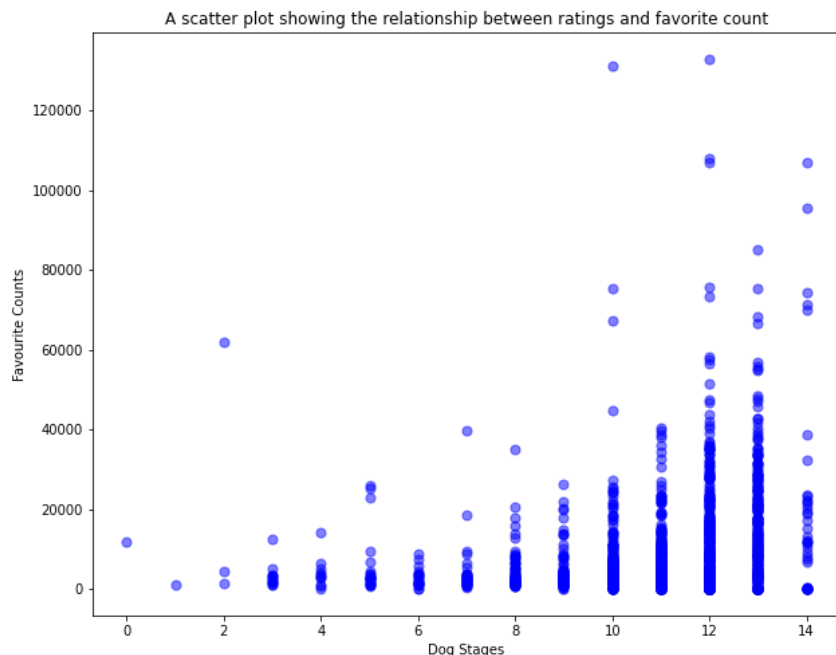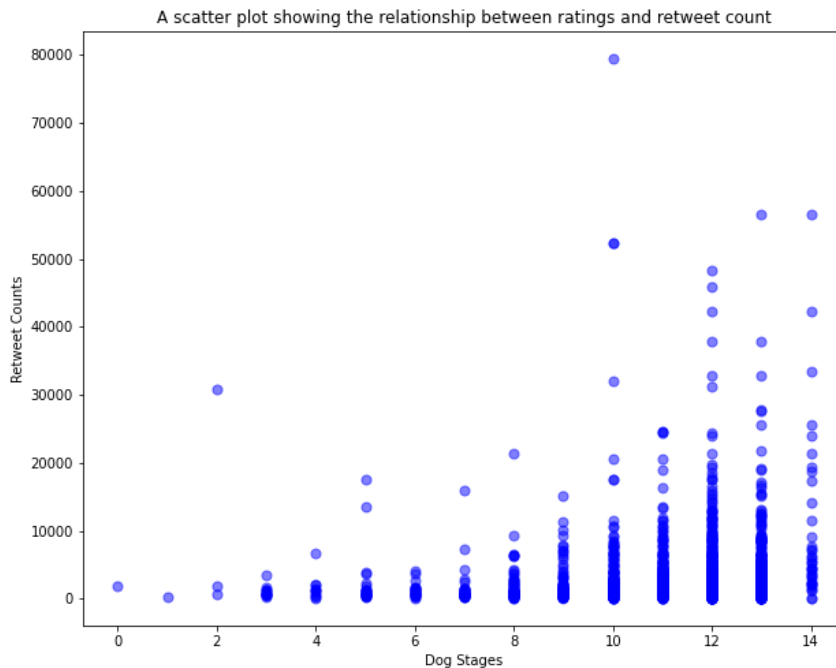
The diagram below shows the mean rating for the various dog stages.
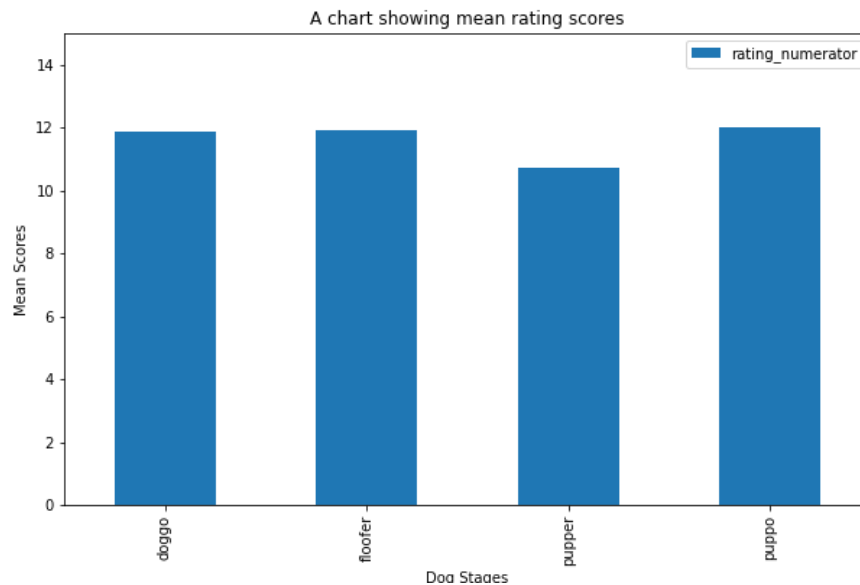
➤ A tweet about each dog includes the favourite count and the number of retweets. The analysis includes the relationship between a dog rating and the favourite and retweet counts.

As indicated in the diagrams below, there is a positive relationship between the ratings that are less than 14 and the favourite count. Thus, as ratings rises to 14, the favourite count also rises.

This is also true for the retweet count. Most tweets with higher ratings have higher retweet counts.

A scatter plot showing the relationship between ratings and retweet count

A scatter plot showing the relationship between ratings and favorite count

➢ Sentiment analysis is a concept used to describe emotions that people hold. It classifies these emotional tones as positive, negative, or neutral. In this analysis, we apply this concept on the tweets (text).

The polarity of most tweets is greater than 0, indicating a positive statement. Thus, most expressions about the dogs are of positive remarks. At the same time, the favourite count of most of these tweets are less 20000.

The subjectivity measures how subjective these tweets are. From the diagram, most of these tweets are concentrated between 0.3 and 0.8. Thus, the subjective are fairly spread, with some highly subjective and others less subjective.

A scatter plot showing the relationship between polarity and favourite count

A scatter plot showing the relationship between subjectivity and favourite count