We Rate Dogs Twitter Data Analysis

By Tal Davidson

Project Overview

WeRateDogs is a popular Twitter account known for rating users' dogs and adding lighthearted comments. Typically, the ratings have a denominator of 10, but the numerators often exceed 10, such as 11/10, 12/10, 13/10, and so on. This reflects the account's fondness for praising dogs as "they're good dogs". However, for dogs that may not be conventionally good-looking or belong to certain breeds, the ratings may be less than 10. With over 4 million followers and widespread media coverage, WeRateDogs is a highly influential account. For this project, I analyzed the Twitter data associated with the account, focusing on data gathering and wrangling skills. Through the analysis, I aimed to gain insights and provide answers to important questions, including exploring correlations among key parameters using a correlation heat map, identifying the most common dog life stage, determining the most frequently assigned ratings, and investigating whether the source of the tweets influences popularity based on the retweet count.

Correlation Heat Map

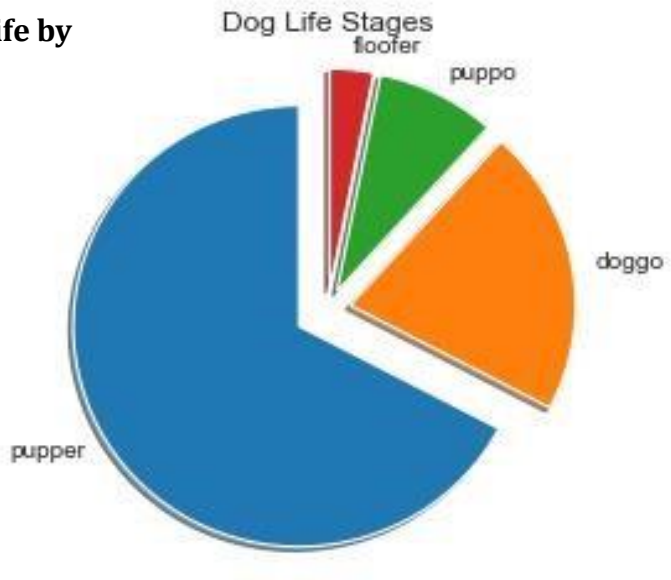
During my analysis, I sought to investigate the relationship between the number of retweets, number of favorites, and ratings. To explore this relationship, I created a correlation heat map. The plot revealed the following insights:

There is a strong correlation (0.9) between the count of favorites and the count of retweets, which aligns with expectations. Generally, when a tweet receives more favorites, it tends to be retweeted more as well.

Surprisingly, there is a low correlation between the numerator ratings and the count of retweets. One might expect that tweets with higher numerator ratings would receive more retweets. However, the data indicates otherwise.

By examining these correlations, I gained a deeper understanding of the relationships between different variables in the dataset and was able to draw insights from the data.

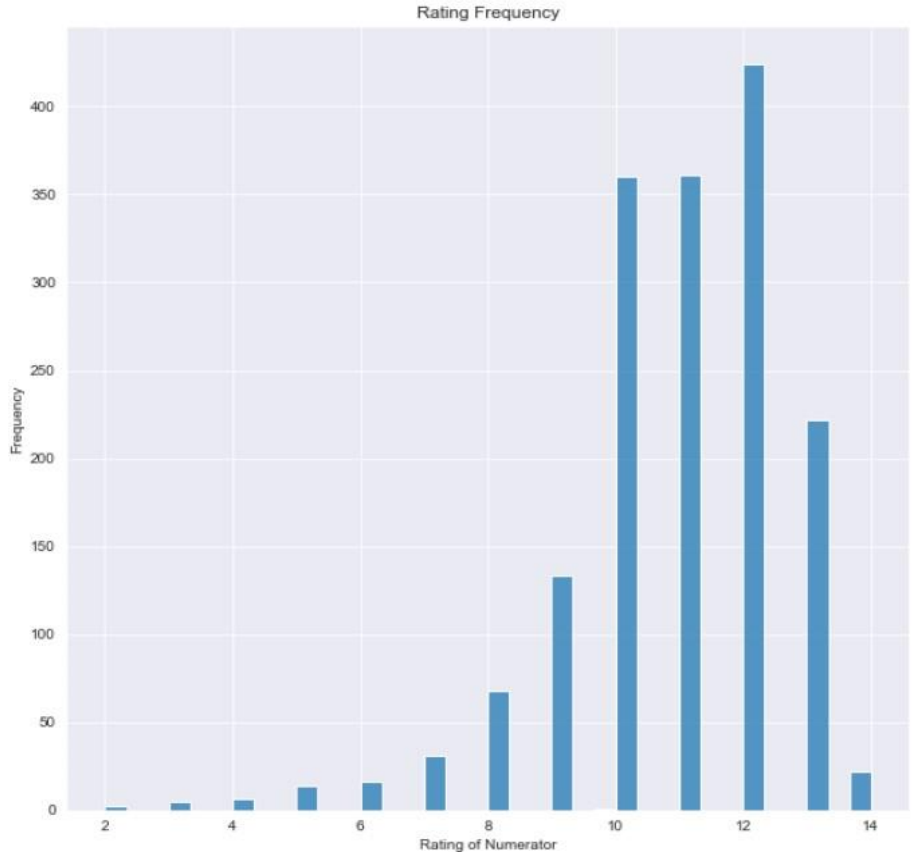
Most common dog life by activity



In my analysis of the master data, I focused on determining the most common dog life stages. To present this finding, I created a pie chart, as shown in the plot above.

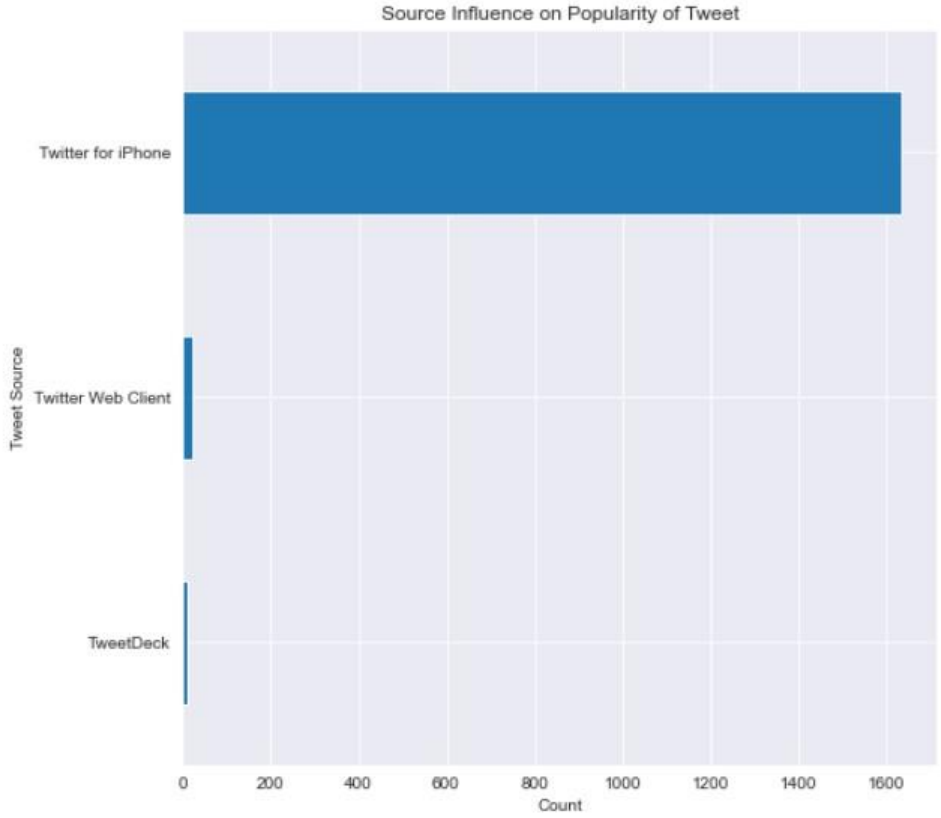
Based on the insights derived from the chart, I concluded that dogs in the "Pupper" stage receive the highest number of tweets and consequently garner the most attention. This finding suggests that dogs in the "Pupper" stage are particularly popular among the Twitter audience.

Most Common Rating



Through my analysis, I discovered that the most common numerator rating for dogs is 12. This finding indicates that a rating of 12 is frequently assigned to dogs in the dataset, suggesting a positive perception of these dogs by the WeRateDogs community.

Does Source influence popularity? (retweet count)



Based on the analysis of the dataset, it was observed that tweets made by iPhone users tend to have higher retweet counts. This finding leads to the conclusion that, in order to maximize engagement, it is advantageous to compose tweets using an iPhone. The data suggests a correlation between the source of the tweet and the level of engagement, with iPhone users experiencing greater retweet activity.

Conclusion

In conclusion, the analysis of this fascinating dataset has been a rewarding and fulfilling experience. It has allowed me to apply the knowledge and skills gained from this course to a real-world scenario that aligns with my interests. The WeRateDogs Twitter page is undoubtedly a fantastic resource for dog lovers, and I highly recommend visiting their page to explore the delightful content they offer. Additionally, don't forget to utilize their rating and retweet functions to actively engage with the community.