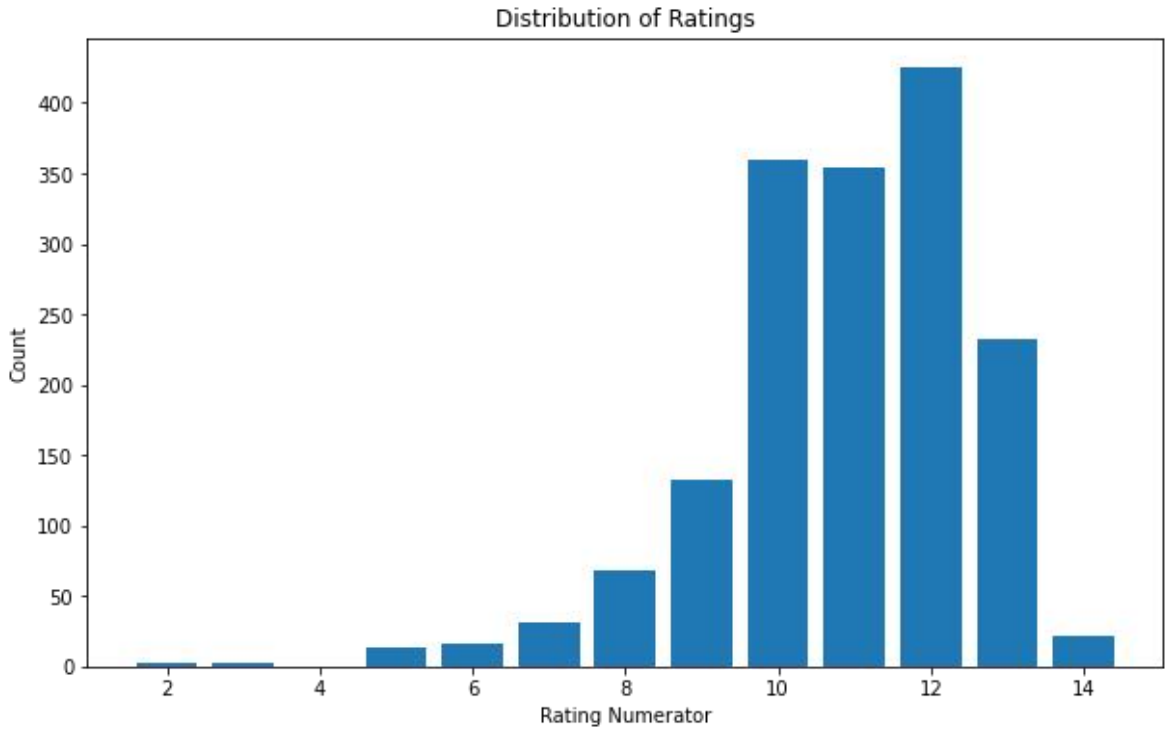
Project: **Wrangling and Analyze Data**
Student: **Mehrol Bazarov**

Udacity 4th project
ACT REPORT
(at least 250 words) - mine (425 words)

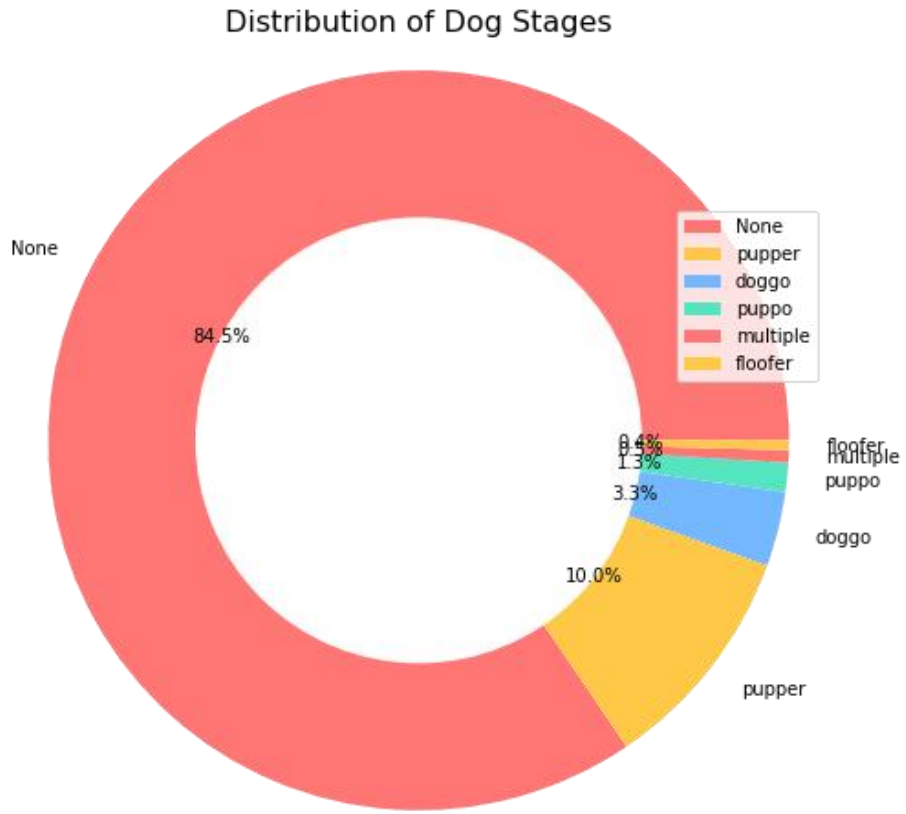
In this report, we will communicate the insights derived from the analysis of the WeRateDogs dataset. The analysis focused on visualizations that provided valuable information about the distribution of ratings, dog stages, popular dog breeds, user engagement by day of the week, and rating trends over time.

The **first visualization** examined the distribution of rating numerators in the dataset. It revealed that the ratings generally followed a normal distribution centered around 12. The majority of ratings fell between 10 and 13, with the highest concentration around 12. This suggests that users tend to give relatively high ratings to the posted dogs.

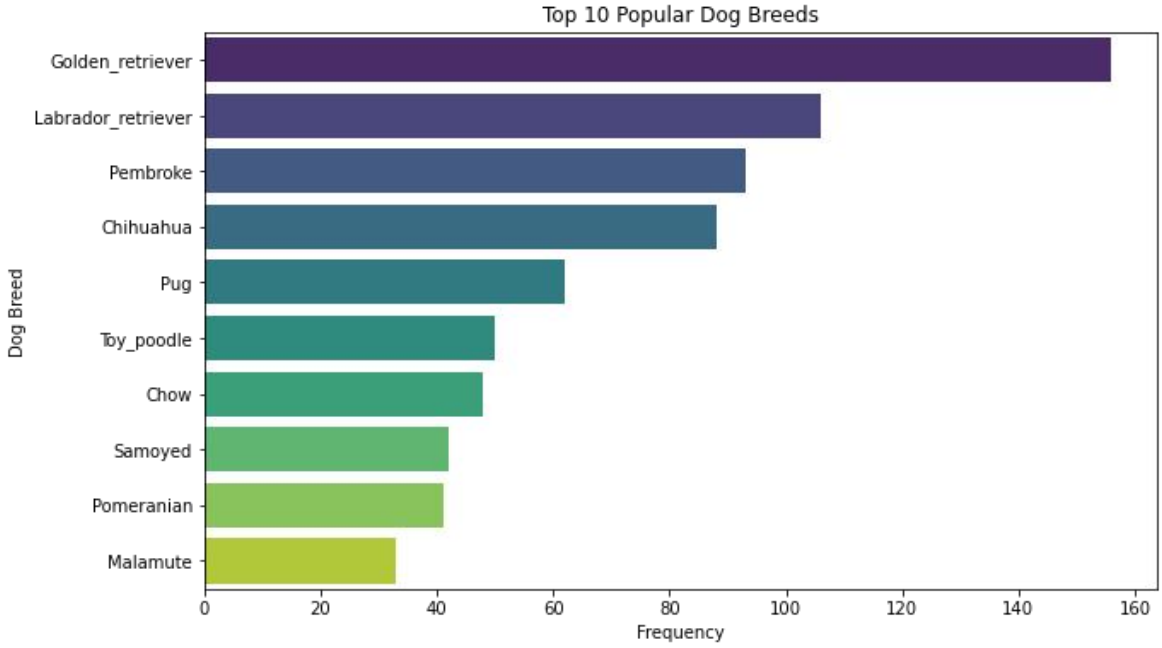


Next, we explored the distribution of dog stages in the dataset. Among the specified dog stages, "pupper" was the most common,

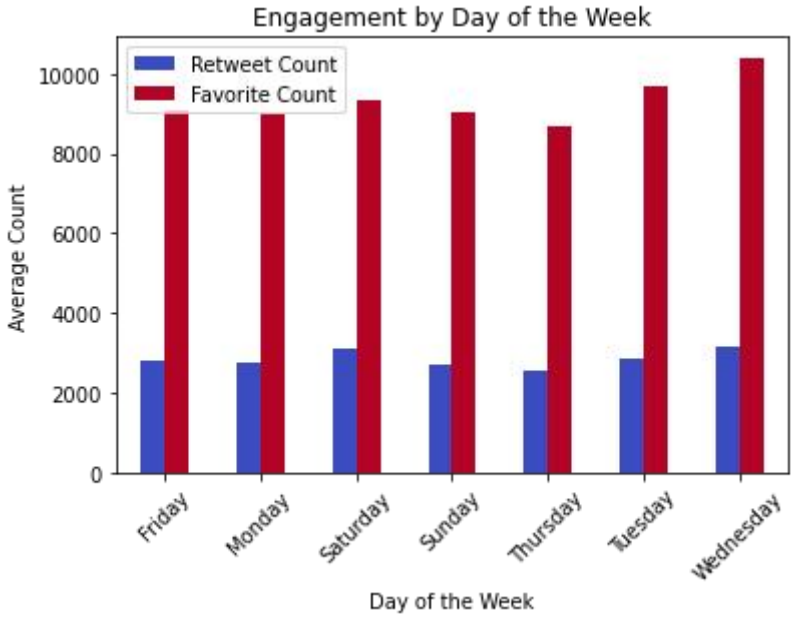
accounting for 10% of the dog posts. "Doggo" and "puppo" were less common, with frequencies of 3.3% and 1.3%, respectively. However, the majority of posts (84.5%) did not specify a particular dog stage. There were also a small number of posts that mentioned "multiple" stages or the "floofer" stage.



We then analyzed the popularity of different dog breeds based on the "breed" column from the image prediction data. The most popular dog breed among the analyzed posts was the Golden Retriever, mentioned in 160 posts. Labrador Retrievers followed closely behind with 110 mentions. On the other hand, the Malamute breed was the least popular, appearing in only 30 posts. These findings shed light on the preferences and interests of users regarding dog breeds.

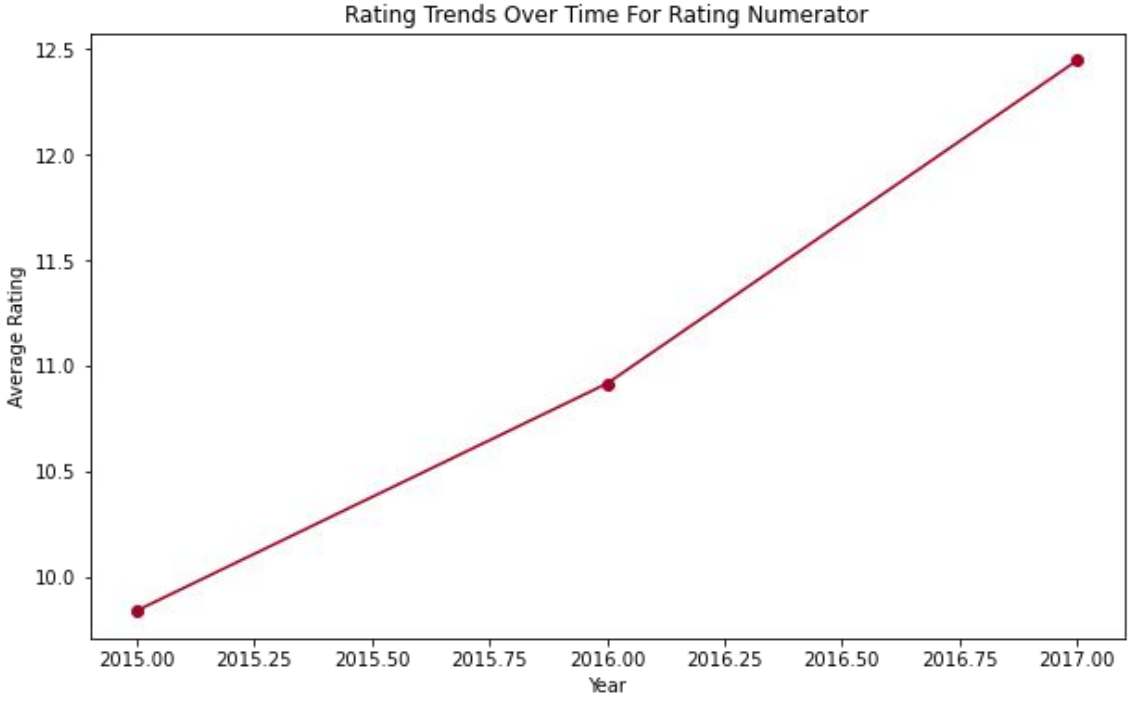


Moving on to user engagement, we explored how retweet and favorite counts varied by day of the week. It was observed that posts made on Wednesday tended to receive more favorite counts. Additionally, Saturday, Wednesday, and Friday showed higher retweet counts compared to other days. These findings suggest that certain days of the week may generate higher engagement and interaction among users.



Lastly, we examined the rating trends over time by grouping the dataset by year or month. The analysis revealed an increasing trend in average ratings over the years. In 2015, the average rating was 9.8,

while in 2017, it had increased to 12.4. This indicates that users have been giving higher ratings to the dogs posted on the platform over time, reflecting a positive sentiment and appreciation for the content.



In conclusion, the analysis of the WeRateDogs dataset provided valuable insights into the distribution of ratings, dog stages, popular dog breeds, user engagement by day of the week, and rating trends over time. These findings contribute to a better understanding of user behavior and preferences on the platform.