Act report – TAIWO, Ridwan

This report details the insights that were gathered from analyzing the cleaned dataset for the "WeRateDogs" project.

Five questions were asked and answered in this report. The questions are highlighted below.

1. Which dog type is most liked in twitter's space?

2. Which dog type is mostly tweeted about?

3. Which dog type has the highest rating?

4. According to the best prediction algorithm, which dog breed received the highest rating?

5. What is the correlation between retweet_count & favorites and rating & favorites?

Insight 1:

The analysis shows that there are 5 types of dogs in the dataset: "None," "Pupper," "Doggo," "Puppo," and "Floofer". The "None" type refers to dogs with an unidentifiable type. Hence, they were removed from the dog-type analysis. The result of the analysis is shown in Figure 1. This result indicates that "Puppo" is the most liked dog, while the least liked is "Pupper."
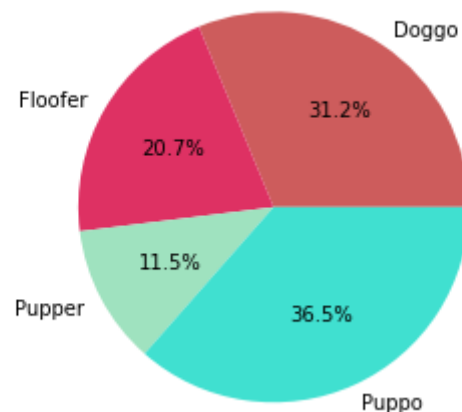

Figure 1: Types of dogs according to their average number of likes

Insight 2:

According to Table 1 and Figure 2, it is evident that "Pupper" is the dog that has the highest number of tweets. Despite the fact that "Pupper" has the most tweets, it is responsible for the least number of likes. This may indicate that many people on Twitter have less admiration for "Pupper."

Table 1: Dog types according to their number of tweets

Dog types	Number of tweets
Pupper	209
Doggo	62
Puppo	23
Floofer	8

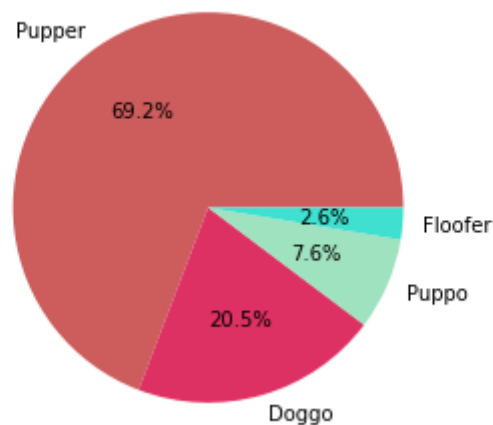Distribution of dog types according to their number of tweets

Figure 2: Types of dogs according to their number of tweets

Insight 3:

The third insight is about the dog type with the highest number of ratings. The analysis, which is visualized in Figure 3, indicates that "Puppo" has the highest number of ratings, while "Pupper" has the least rating. This result conforms with that of the first insight, indicating that "Puppo" is a dog-type with desired qualities.
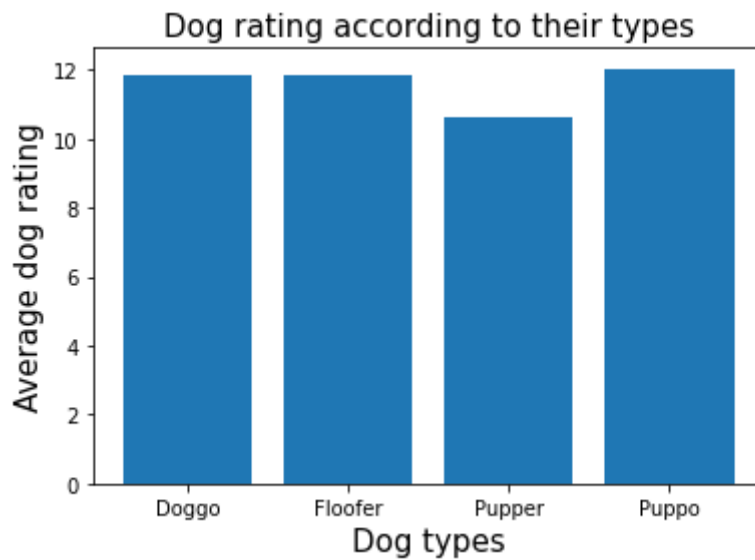
Figure 3: Rating of dog types

Insight 4:

As per the best prediction algorithm, the dog breeds with the highest ratings are "pedestal," "military_uniform," "limousine," "pole," "prison," and laptop, as they all have a mean rating of 14/10. The top 30 dog breeds are presented in Figure 4.
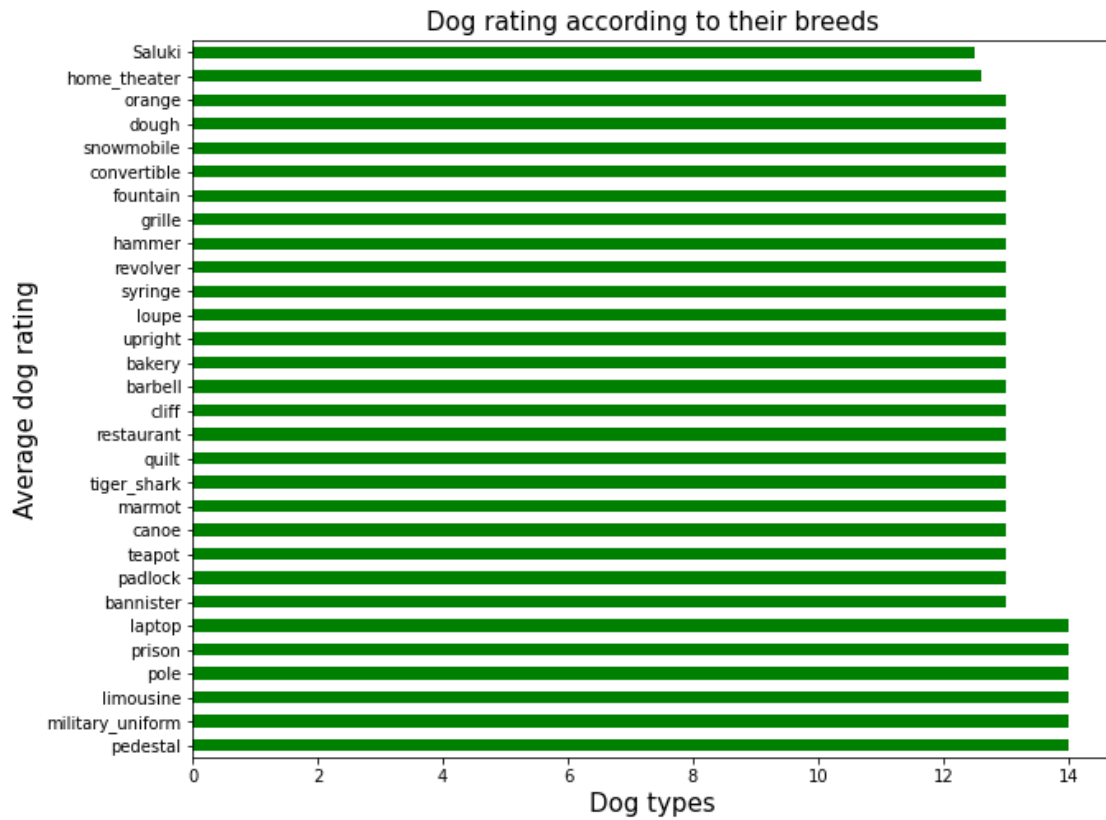
Figure 4: Top 30 dog breeds according to the best algorithm

Insight 5:

According to the result (see Figure 5), there is a high positive correlation between retweets_counts and favorites, as the correlation coefficient is 0.93. This is intuitively true as dog types with many favorites (i.e., likes) have the tendency to be retweeted many times. Furthermore, the result shows that there is a positive relationship between dog rating and favorites.
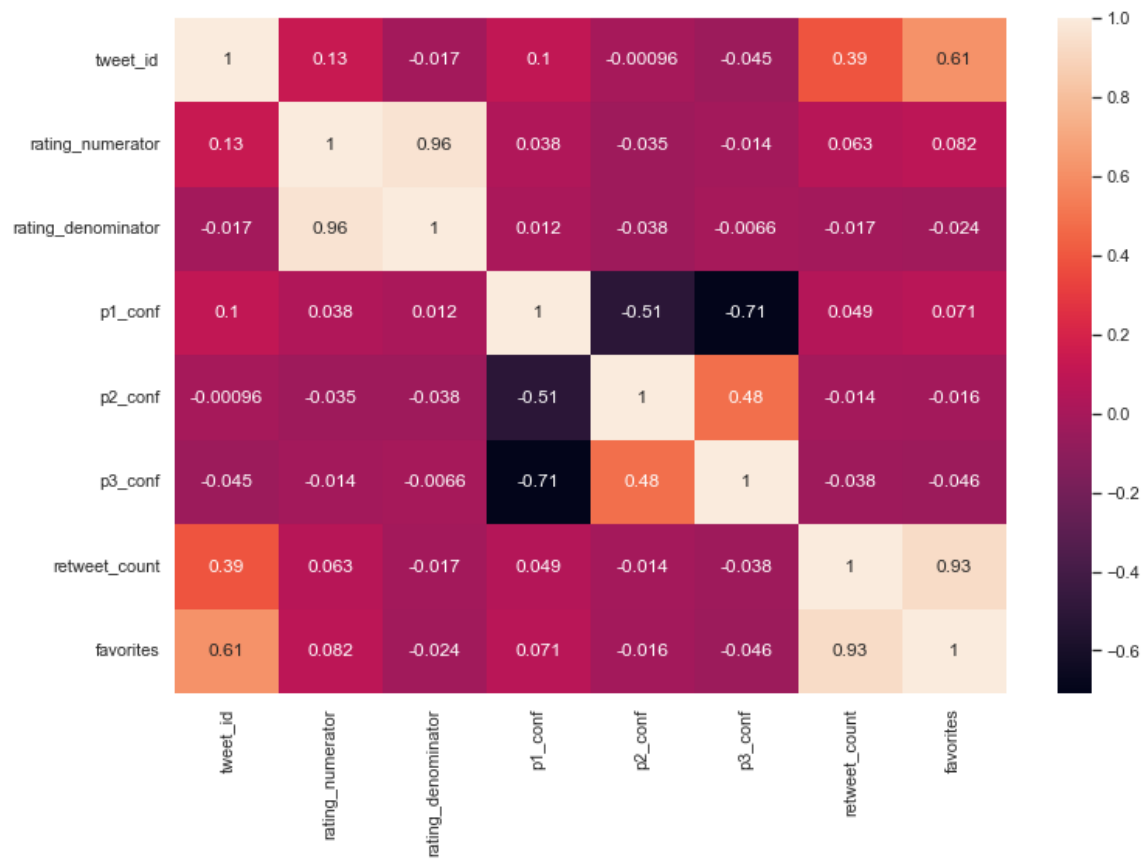
Figure 5: Correlation maps between the numerical variables