Act report

Kexin Yao

In this project, after wrangle the data of WeRateDogs tweet, analysis and visualization based on the cleaned data was conducted.

The analysis questions of the data is following:

1. What is the most popular name(s) for dogs?
2. Which breed of dogs has the highest rating?(top10)
3. What stage of dogs do people like most?
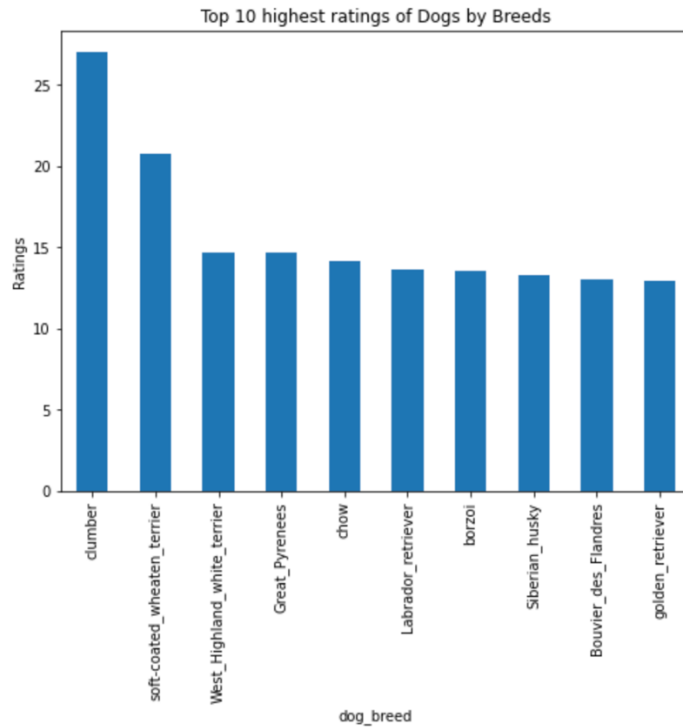4. What is the relation between rating and retweet counts?

1. What is the most popular name(s) for dogs?

   By using the pandas function value_counts, we can found the top 10 popular names for dogs is in the table below. As we can see, the most popular name for dogs is **Charlie,** which have counts of 11.

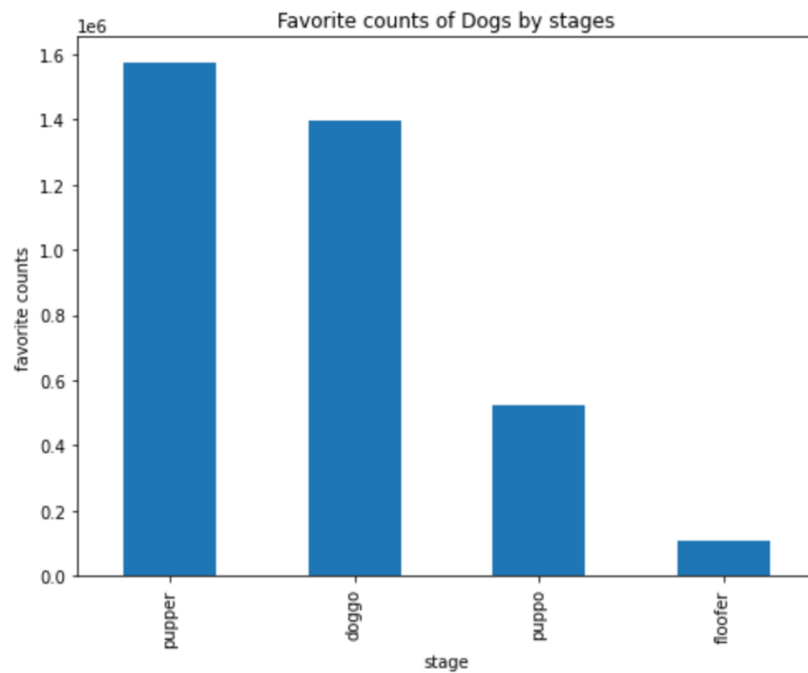| Dog Name	Counts
Charlie	11
Cooper	10
Oliver	10
Lucy	10
Tucker	9
Penny	9
Winston	8
Sadie	8
Toby	7
Lola	7

2. Which breed of dogs has the highest rating? (top10)

   Creating a new columns about the breeds of dogs and averaging the ratings for each breed, and plot the bar chart. From the bar chart we can see the highest rating dog is dumber, and the second highest one soft-coated_wheaten_terrier. Although we plot the top 10 highest rating breeds of dogs, the ratings through third highest to tenth highest are quite propinquity

Top 10 highest ratings of Dogs by Breeds

3. What stage of dogs do people like most?

Grouping the dog's stage and get the sum of their favorite counts.The results are visualizing by the bar chart below. Dogs belong to the pupper stage have the most favorite counts, and the second is doggo. Dogs with floofer stage have the least favorite counts.

Favorite counts of Dogs by stages

4.What is the relation between rating and retweet counts?

To investigate the relationship between the rating and the retweet counts, a scatter plot was created. From the scatter plot following, there is no obvious relationship between the rating and retweet counts.

Relation between Dog Ratings and Retweet Counts