

Act Report for WeRateDog Data:

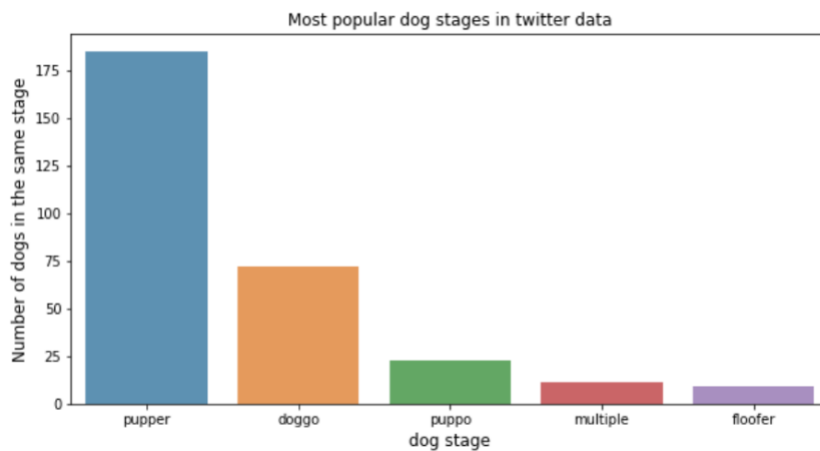
Data Wrangle:

8 quality issues and 2 tidiness issues are fixed before proceeding to data visualization and analysis.

Data Visualization and Data Analysis:

1. What is the most popular dog stage?

There are four types of the dog stages in the dataset, I am interested to know which dog stage is the most popular one. Therefore, I used Panda's `value_counts()` and seaborn to plot out the bar graph of the number of occurrence of dog in each dog stage.



Based on the graph, pupper is the most popular dog stage.

2. What is the most popular breed of dog?

In order to know the most popular breed, I used `value_counts()` on 'p1' column, which is the top prediction of breed in the dataset. Then I changed column name of 'p1' to 'type' to make the new dataframe easier to understand. According to the dataframe, Golden Retriever is the most popular breed with 139 counts. Therefore, the most popular breed of dog is Golden Retriever.

	type	counts
0	golden_retriever	139
1	Labrador_retriever	95
2	Pembroke	88
3	Chihuahua	79
4	pug	54

3. What breed has the highest rating among top 20 most popular breeds?

As mentioned previously, we have the dataframe of most popular breeds (by counts). Now we use another dataframe created before, which is the 'rating' dataframe. This one includes all the rating of different breeds. I have joined the 'rating' table with 'valueCounts' table to get both the rating and counts into one dataframe. Next, I select the top 20 most breeds with most counts, and rank the rating from in descending order.

	type	counts	dogRating
0	golden_retriever	139	11.755162
6	Samoyed	40	11.700000
5	chow	41	11.609756
2	Pembroke	88	11.443182
10	cocker_spaniel	27	11.296296
11	French_bulldog	26	11.269231
8	Pomeranian	38	11.151316
1	Labrador_retriever	95	11.126316
7	toy_poodle	38	11.105263
9	malamute	29	10.896552
12	Chesapeake_Bay_retri...	23	10.782609
3	Chihuahua	79	10.708861
13	miniature_pinscher	22	10.590909
4	pug	54	10.240741

As a result, Golden Retriever has the highest rating among top 20 most popular breeds, with a top score of 11.76.