

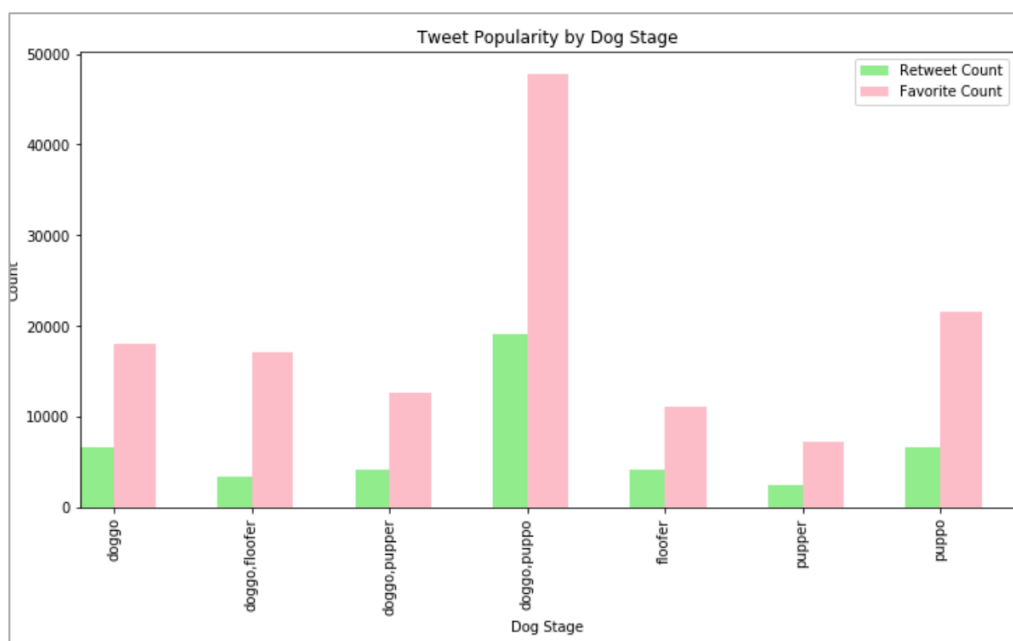
Data Analysis for WeRateDogs Twitter Page

Data analysis here was done after spotting insight ideas that I can explore in the assessment process. I explored few ideas then I picked three for the project. The most interesting parts in my data that I can play with are dog stage, rt and fav counts, rating and god breeds. First, I chose to see what dog stage gets the most rt and fav so I used:

```
df_master.groupby(['stage']).rt_count.mean()
df_master.groupby(['stage']).fav_count.mean()
```

To get the average of all rt and fav grouped by dog stage. I understand the number of tweets that have one stage can be greater than another stage, which can effecting this result greatly, but I chose to not look at his now. The result is that when puppos and doggos are together it gets the most favorites and retweets making it the most popular. The least average of rt and fav is for puppers.

In the three insights that I chose, this idea would be much easier if gets visualized, so I chose to use bar graph to display rt and fav count for each stage:



For my second insight, I chose to see the most common dog breed in the WeRateDogs twitter page. I used value_counts for the top 3 dog breeds and found that retriever dog breeds are the most common.

The third insight is the ratings for each stage. The average rating grouped by dog stage showed here also, the average rating for puppo and doggo is higher than the other dog stages. Puppies get lower rating same as the first insight.