

Ahmed Mousa

s-ahmed.mosa@zewailcity.edu.eg

Udacity data wrangling project visualization and insight report

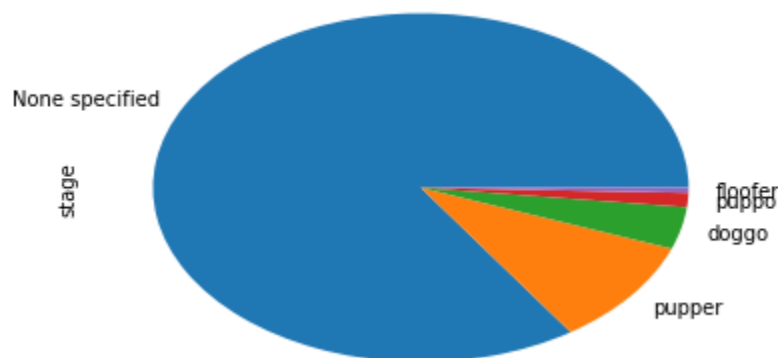
The insights discussed in this report are related to retweets and favorite counts. Our goal is to see which stages get more retweets and have more favorite counts, and the same for dog counts; we want to know whether tweets featuring more dogs get more retweets and have more favorite counts.

Dog stages retweets and favorite counts

Before trying to find which stage has more retweets and favorite counts in the final data frame, we need first to find their distribution to know whether one or more stage overwhelms the rest.

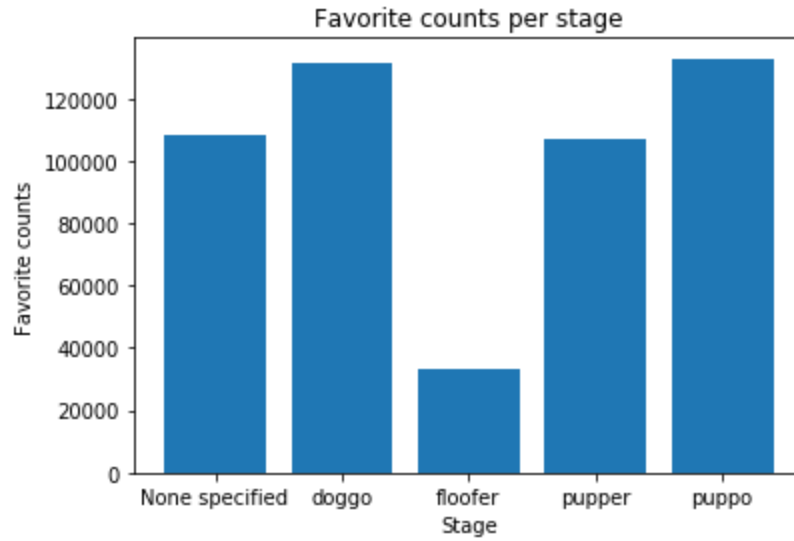
1. Pie chart for numbers of stages

Although most tweets do not specify a stage, the descending order of stages according to the pie chart shown is: pupper > doggo > puppo > floofer. Hence, we should expect higher retweets and favorites for pupper and doggo stages.



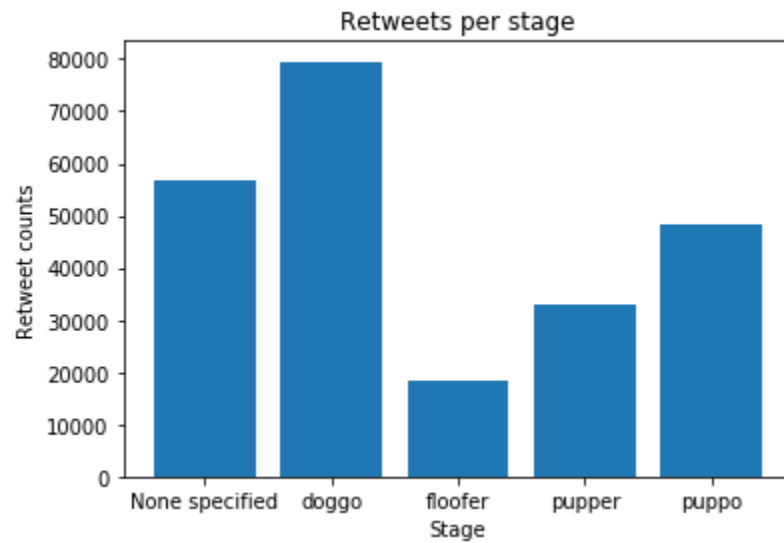
2. Favorite counts per stage

The shown bar chart meets our expectations in regards to floofers having the lowest favorite counts. However, unexpectedly, puppos and doggos are approximately equal, and both are higher than doggos contrary to the presence order.



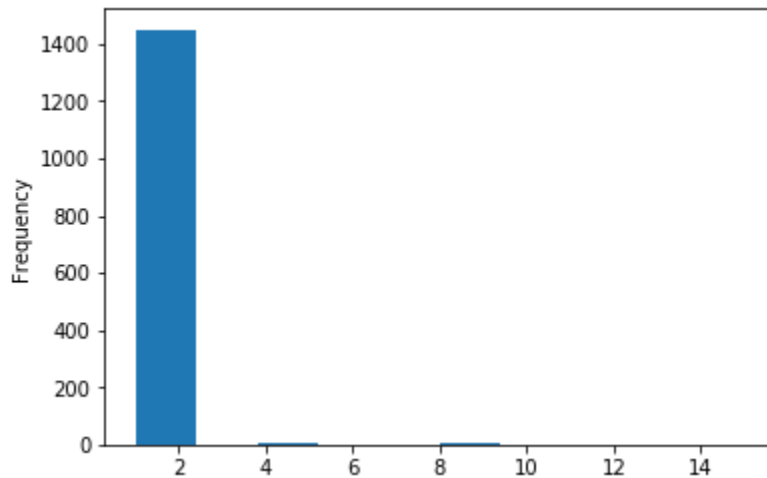
3. Retweets counts per stage

The same order is present as in the previous favorite counts graph. However, doggos here have more retweets than puppos with a huge difference.



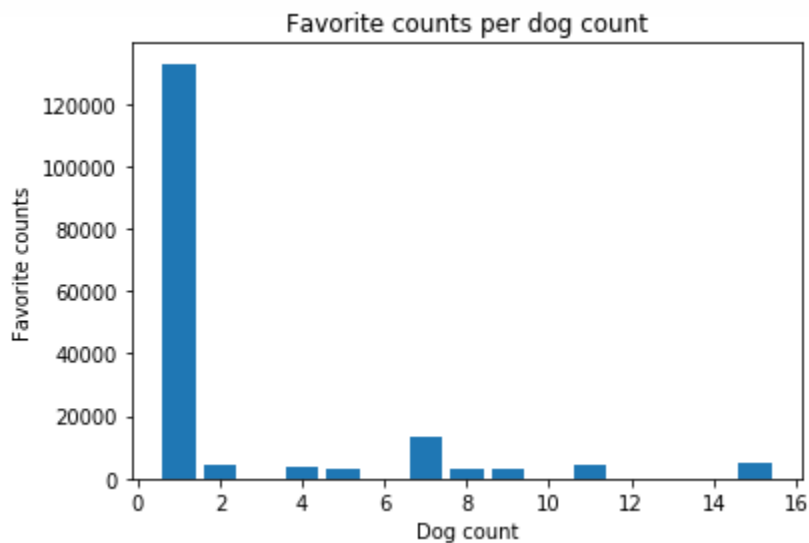
Dog counts retweets and favorite counts

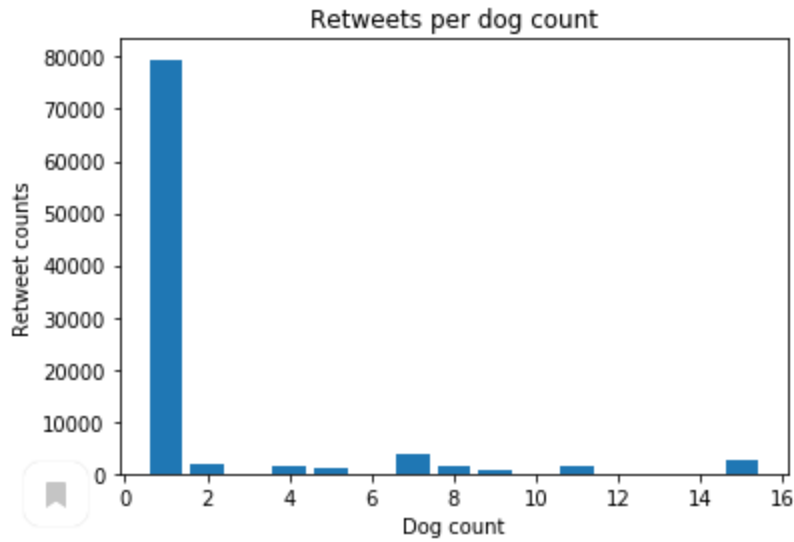
As before, we need to check the counts to know if one count is significantly more present over others. As shown in the next histogram, the counts of 1 is significantly more present. Thus, we would expect more retweets and favorites for them.



1. Favorite counts and retweets per dog count

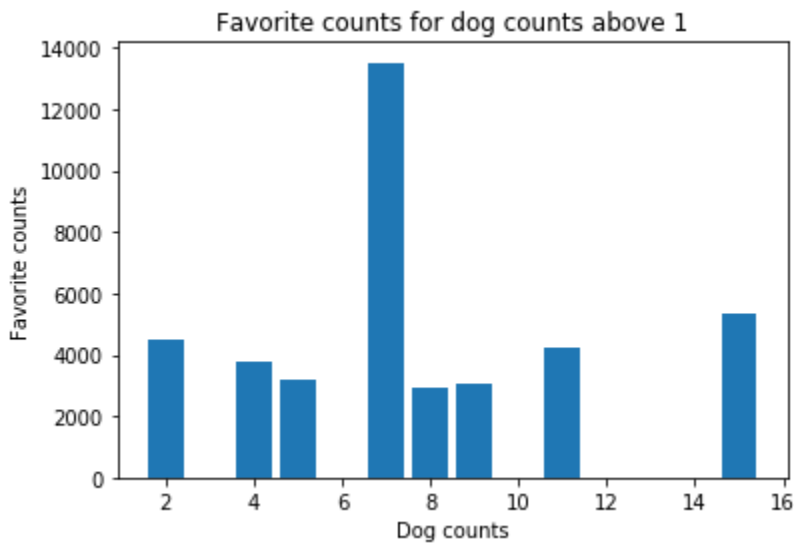
The next two graphs for favorite counts and retweets show that there is no comparison between the count of 1 and all other counts.

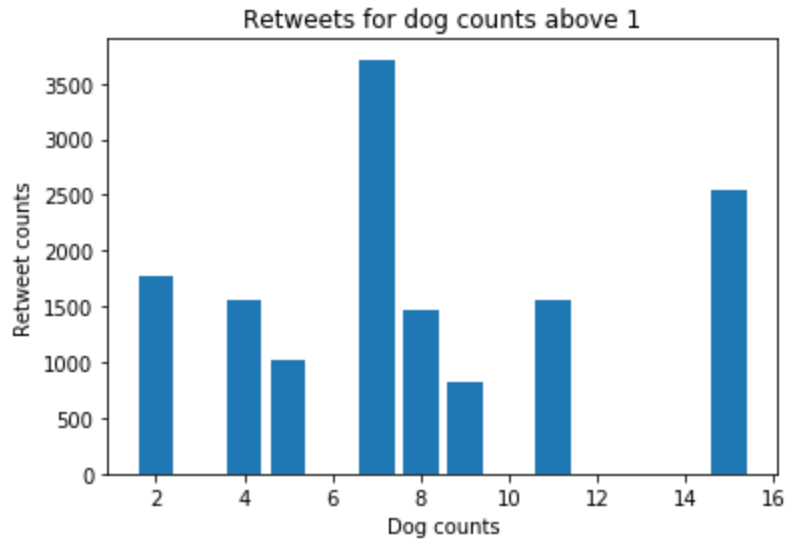




2. Favorite counts and retweets for dog counts above 1

Since we want a clearer insight, we shall exclude 1 and compare favorite counts and retweets for counts above 1. As shown in the next two graphs, tweets featuring 7 dogs receive more retweets and favorite counts above others.





Also, we could calculate the means for favorite counts and retweets for the count of 1 to see how these counts compare. For favorite counts, it is approximately 9371. For retweets, it is about 2811. As we can see, only the count of 7 surpasses the means in both graphs, with the count of 15 close to the mean in the retweets graph.

Conclusion

So, in conclusion, more dogs do not necessarily get more retweets and favorite counts; and doggos and puppos receive higher retweets and favorite counts.