act_report.pdf

September 20, 2018

It looks like the dog data can best be broken down by investigating how the dog type attribute when populated affects things. If we look at the first graph (count by dog type) we see that the pupper dog type shows up in the data set the most, besides NaN values. The next graph looks at the relationship between the various dog types and how many favorites they got cumulatively. It looks like the pupper dogs, which where the most represented in the data set, have a higher favorite count than the other types besides the NaN values. This is unsurprising as the pupper dog type shows up more than the others. In fact if you look at the favorite count chart, it looks like the dog types have just as many favorites as they appear in the data. Which means that people generally favorite the different kind of dog types irrespective of their actual type. This means that the dog type variable probably won't provide much insight to us on which types of dogs get more favorites or any other variable. This is also most likely because the dog type variable has been relatively arbitrarily arrived at based upon whoever created the data set in the first place.

The second plot provides us with a look at the retweet and favorite counts for this twitter account over time. It looks like both have slowed down as the time has gone on.