Act report

As a student at Udacity enrolled to NanoDegree professional track of data analysis, I have the ground skills to be able to analyze and drive meaningful insights and interactive visualizations from the data I just wrangled. After wrangling the data of We Rate Dogs which makes tweets including a dog's image and rate with their different but interesting rating system. Further, With the attached photo we could inferred the breeds of the dog   through neural network constructed by other colleges. lastly, by data queried from tweets itself about the number of the tweet has been favorite and retweeted. Thus, we derived 3 insights almost collecting those efforts altogether and also 3 visualizations.

Firstly, insights

1. we can give some insights about the stage of the dog. from the tweets that have available stage name, more than half of their dogs are pupper. In deed , (**65**%) of the rated dogs are in the "pupper" stage while "doggo" dogs proportion is (**21.5**%). Floofer is very rare amounted to only (**2.7**%) analogous to the tweets which has more the 1 dog reached only (**2.3**%) and has both doggo and pupper dogs. This interesting as if the page(we rate dogs) aims to plan cooperation with other pages uses this target segment of dogs, it has to chose the pages targeting (pupper) mainly.
2. the correlation between 'rating_numerator', 'favorite_count','retweet_count'., 'favorite_count','retweet_count' that gathering 2 files data together worth analyzing.

variable	rating_numerator	favorite_count	retweet_count
rating_numerator	1.000000	0.424461	0.314499
favorite_count	0.424461	1.000000	0.913173
retweet_count	0.314499	0.913173	1.000000

As expected, the higher rating_numerator, the higher favorite_count and retweet_count infered by the positive sign. However the relationship is not very strong between rating_numerator and favorite_count or between rating_numerator and retwee_count(below 0.5), supposingly high rated dogs don't receive much cheers!. But the correlation between favorite_count and retweet_count is strong (0.9), normally intuitive.

3. finding the most common dog's breeds of the dogs which lie in the 25% top ratings illustrated by the 3 algorthims. Here, we have used also 2 data frames to infer about those dogs who got excellence ratings which amounted to 200 with their breeds configured by the colleges.
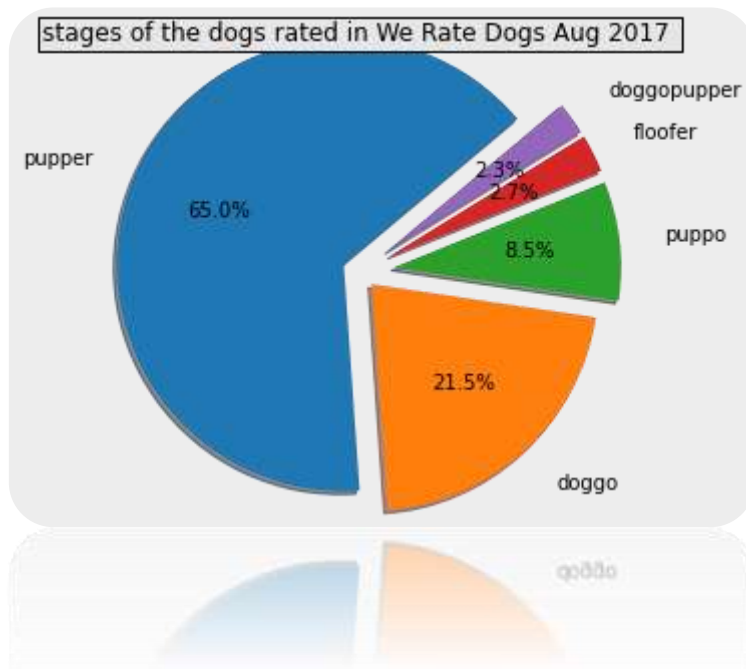Most common breeds are 4 as follows:
   a. **Golden_Retriever**, which exists by 30%, 11% or 7% according to algorithm 1,2 and 3 respectively.
   b. **Labrador_Retriever**, which exists by 22%, 13% or 12% according to algorithm 3,2 and 1 respectively.
   c. **Chihuahua** , which exists by 10% or 9% or according to algorithm 1 and 3 respectively.
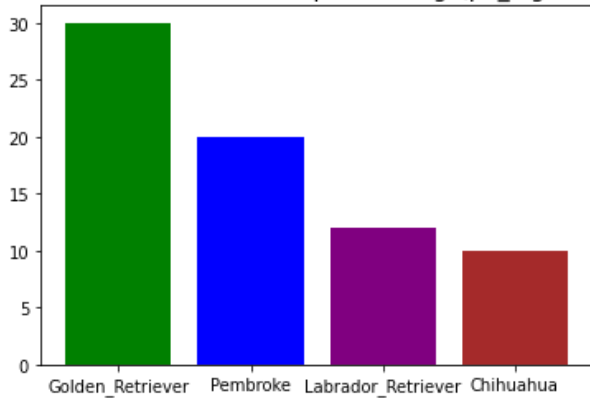
Secondly, visualizations

We illustrated 1 and 3 in the above insight to visuals to interact with them comprehensively.

1- The first figure as mentioned earlier demonstrated the proportions of the stage classes that were reported in the tweets excluding the tweets that has not mentioned the dig's stage in a pie chart.
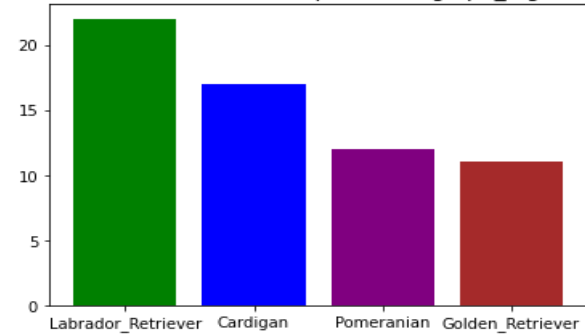
stages of the dogs rated in We Rate Dogs Aug 2017

2- The following bar charts shows the most common breeds of the dogs that are top rated according to its configuration algorithm.