Act Report

The Dog or domestic Dog is a domesticated descendant of the wolf. There are different species of the dog. Dog breeds vary widely in shape, size and colour.

Over the years dogs have been rated by humans according to their own personal preference, these dog ratings in this case have been collected, Ratings was sent out through tweets which was scrapped using twitter API and was downloaded, The data then wrangled and cleaned for analysis purpose.

After the overall wrangling and cleaning of the master data, its about time to make some analysis and visualisations.

First thing first, importing of necessary libraries, in this case pandas as pd and numpy as np, then load the master dataset that have been wrangled, cleaned and combined into the notebook.

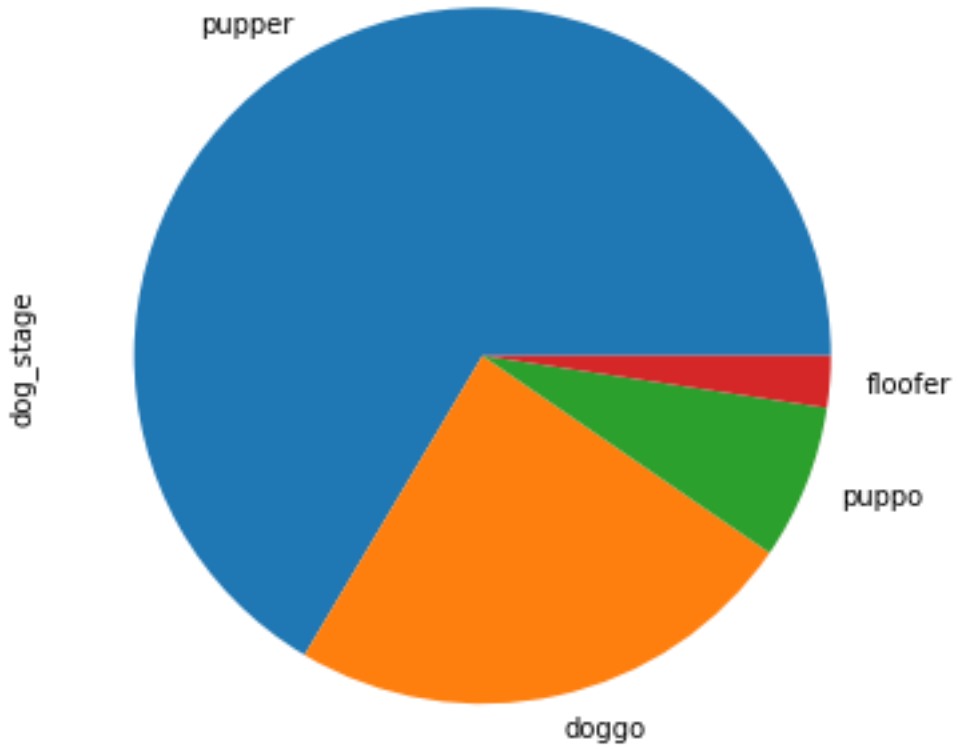
Since the dataset contains rating numerator and rating numerator, I then proceeded to create another column which is the division of the numerator by the denominator with name ratings.

Below are my questions and end findings from the wrangling and visualisation of the dataset.

1. The first question that came into my mind was to find out which dog has the highest number of ratings, then from my analysis we get to understand Dog name Aticus has the highest rating of 177.6

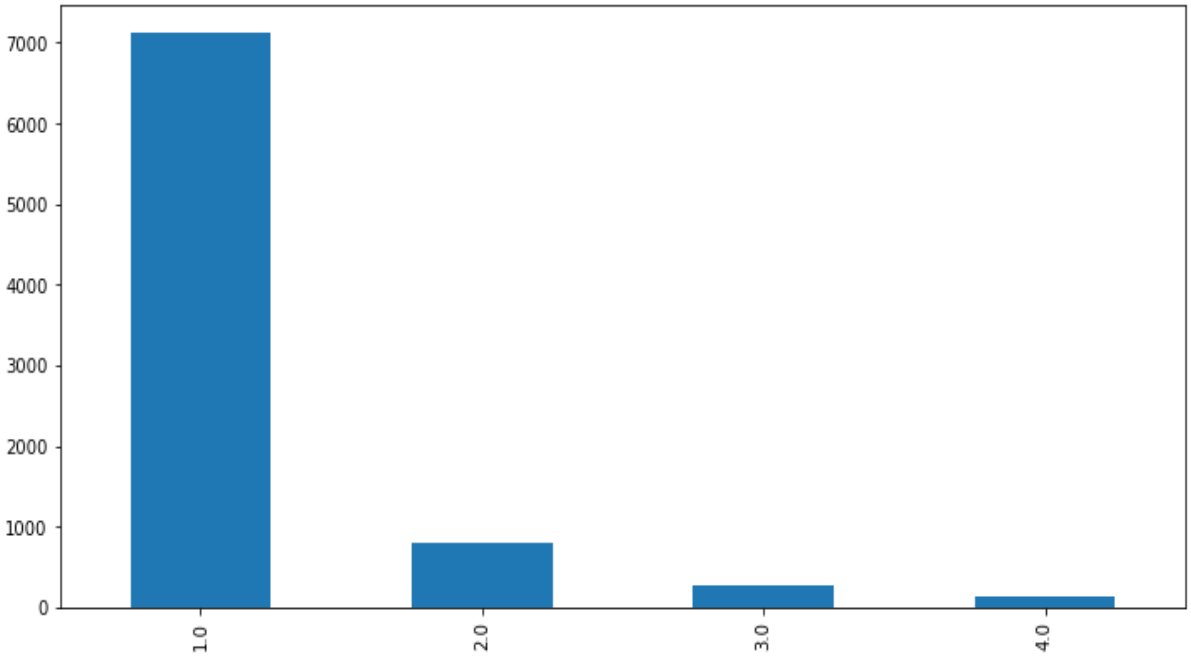
2. Furthermore, we wanted to investigate the average retweets per tweet from the dataset, by the analysis we got to understand that the average retweet count of the tweets is 2977.

3. Question: Finding out which of the dog stages could be the most occurring or simply put the dominant one in the dataset?



Findings: The result from this analysis shows that dog stage Pupper is the dominant dog stage in the dataset.

4. Question: Lastly considering the img_num column, similarly we checked for the dominant image number.



Findings: The result of the analysis dhows that image_ number 1 is the most dominant image number from the dataset.