Report: act_report

- Create a **250-word-minimum written report** called "act_report.pdf" or "act_report.html" that communicates the insights and displays the visualization(s) produced from your wrangled data. This is to be framed as an external document, like a blog post or magazine article, for example.

Loading of saved file

The saved **twitter_archive_master.csv** was read to a dataframe using pandas read_csv() function and by passing tweet_id and timestamp column datatype. The file was read to **master_df.** The **master_df** was opened and irrelevant columns to my analysis were dropped using drop() function

Insights

Using describe() function and by visualising the data, below insights were noted

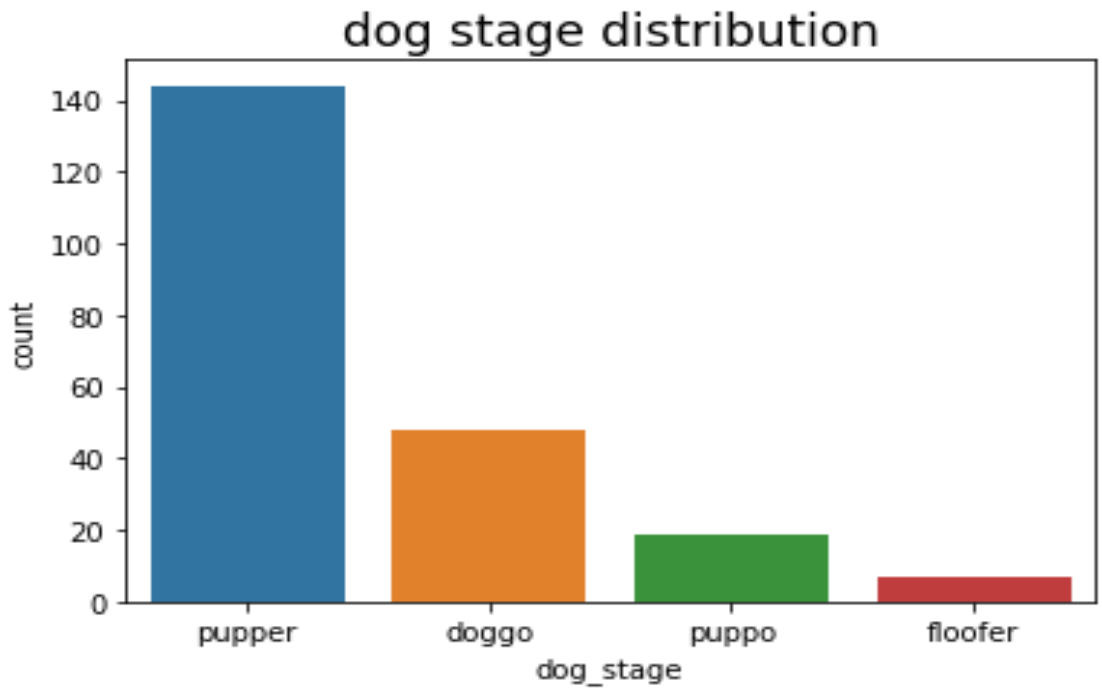
- pupper dog_stage is frequently tweeted
- increase in retweet count led to increase in favorite count
- the frequent number of images in a tweet is 1
- the maximum number of favorite count is 132810

Analysing and Visualizing Data

From my insight, 2 questions were asked and analysis with visualization were carried out.

Question 1: what dog_stage is frequently tweeted?

The frequently tweeted dog stage in all the tweets gathered after cleaning is **PUPPO.** And the least is **FLOOFER**. This shows that the people that tweets were having more PUPPO as their dogs and few were having FLOOFER.



Question 2: What is the relationship between retweet count, favorite count, image number, and predictions of dog breed?

- favorite_count is strongly positively correlated with retweet_count
- img_num and p1_conf are far less correlated
- favorite_count and p1_conf are negatively correlated
- retweet_count and p1_conf are negatively correlated
- favorite_count and img_num are negatively correlated
- retweet_count and img_num are negatively correlated

It shows that more retweets attract more multiple favorites. Tweets with 1 image is likely to be retweeted and liked than tweets with 4 images.