

I. Insights

I answered three question of dataframe.

- 1) What's type of image is the best rating?
- 2) What's is the dog name most popular in twitter?
- 3) In Which month are there more Tweets?

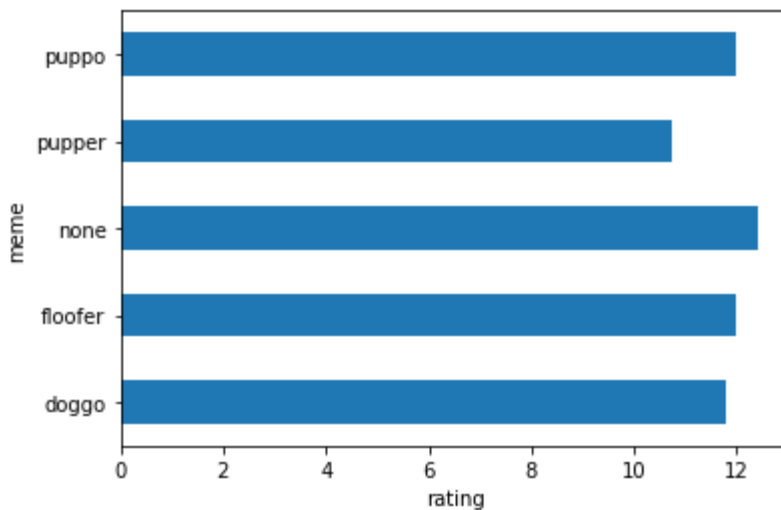
The three insights have a plot that help to visualize the results.

For this work I used the cleaned information that was saved and stored in a csv (df_tweeter_clean.csv)

What's type of image is the best rating?

For answer this question I used the mean of column rating (this column is rating_numerator /rating_denominator after clean rating_denominator).

I plot the results in a bar plot.



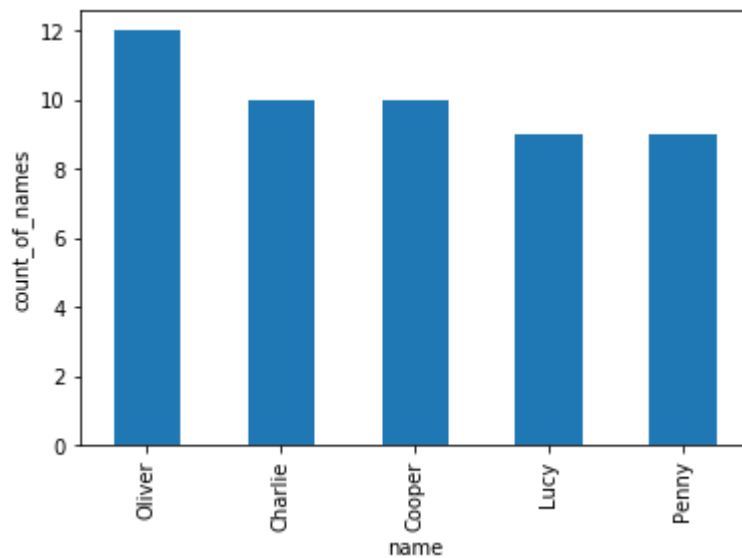
The most popular type of image us puppo. The second one is flooper and the third one is doggo

The most unpopular is pupper.

What's is the dog name most popular in twitter?

For answer this question I count the names of the dog, after of clean this column (in this column there are many word with error transcription , for example names call: a , an , this , my, etc.)

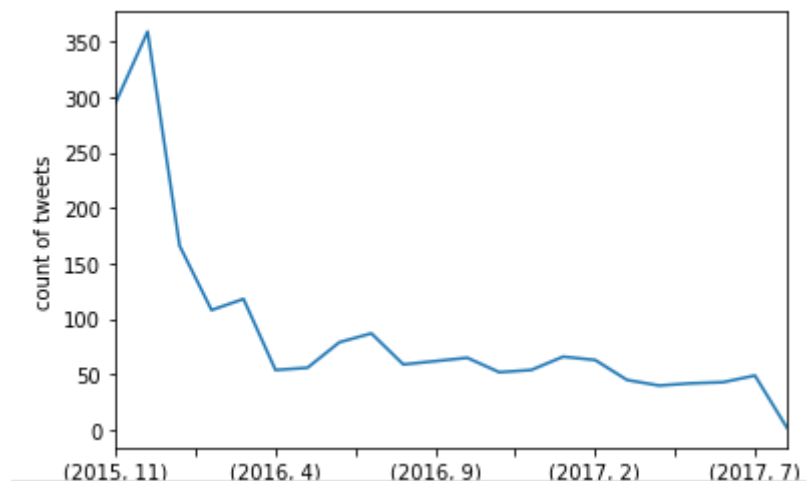
I plot the five more common names of dog.



The most common dog names are Oliver, Charlie, Lucy, Cooper and Lola.

In Which month are there more Tweets?

For answer that question I plot a time series with the numbers of tweets for each month.



In the plot is observed that in December 2015 is the peak of tweets, after that, the numbers of tweet decrease considerably. In fact in the last month of the time series is the minimum number of tweets of the all range of time.