Act report

By making wrangling process to data, I find a lot of insight in data that make decision easy to take; so, I will show these insights to you.

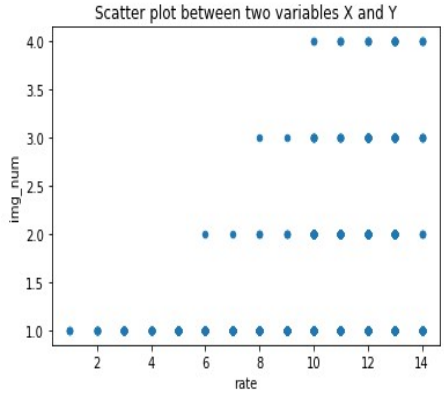
During assess phase I find some insights, it is:

1. In stage of dogs pupper stage is the most frequent value.

Then in final dataframe I do some diagrams that tell me some insights,

I will add screens from
them.

```
In [75]: master_dataframe.plot.scatter(x='rate', y='img_num', title= "Scatter plot between two variables X and Y");
         plot.show(block=True);
```

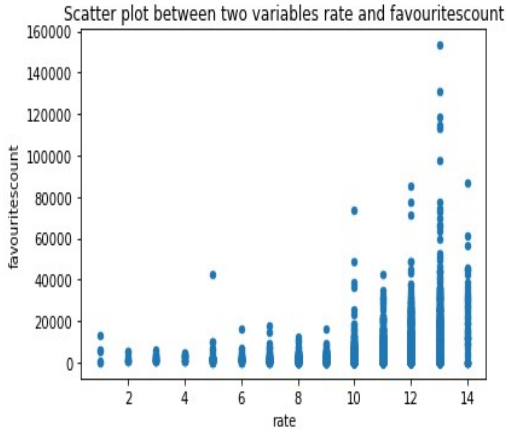


At this figure I make a scatter plot between number of images in the tweet and the rate in text.

In this image you see that when number of images was 1 all levels of rate exist, but when number of images become 2 or 3 or 4, the small rate become not exist.

That tell us that more the number of image increase, more the rate increase too.

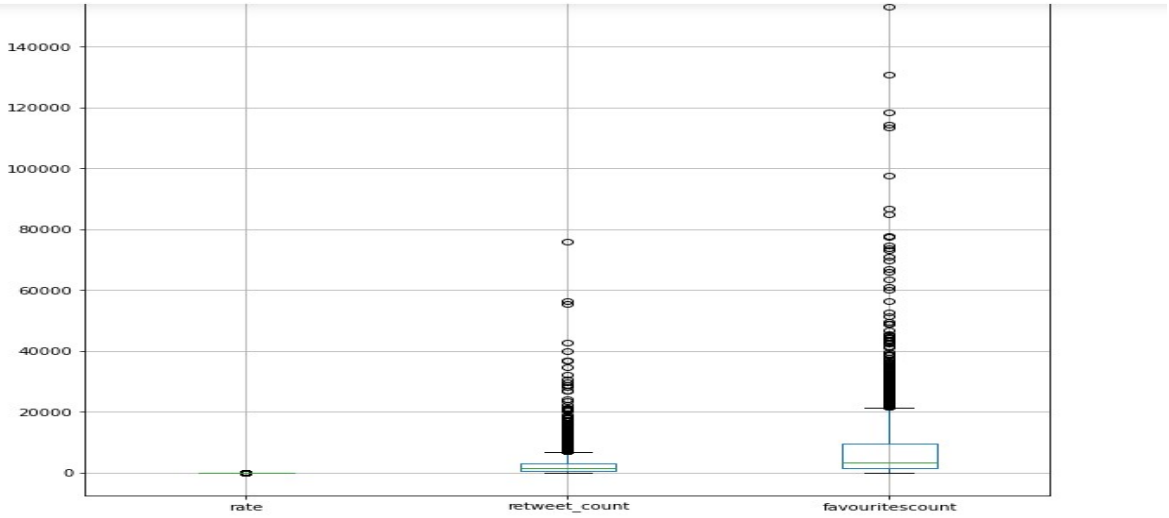
```
In [76]: master_dataframe.plot.scatter(x='rate', y='favouritescount', title= "Scatter plot between two variables rate and favouritescount'
         plot.show(block=True);
```



In this figure I make a scatter plot between favorites count and rate.

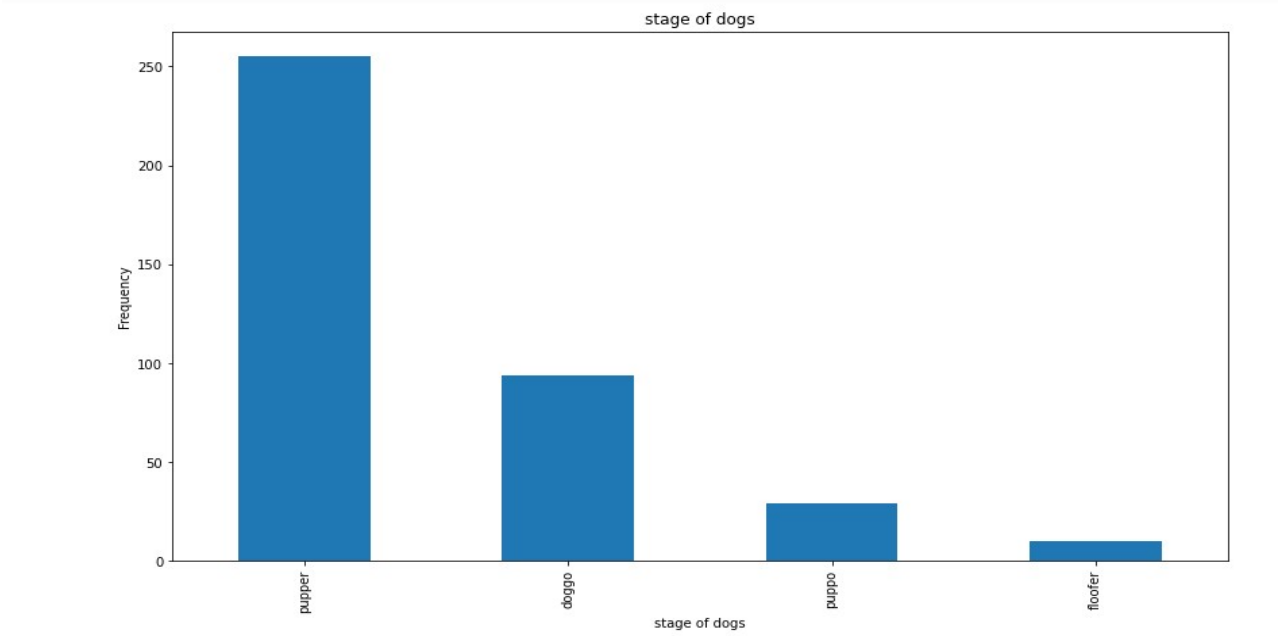
So, you see that if the rate of dog is 10 or 14 the number of favorites to this tweet become great.

By the way, that tell us that the more rate increase, the more favorites count increase.



By this figure and the mean of rate and retweet count and favorite count.

We find that rate not have a outliers but retweet count and favorites count have a lot of outliers, that mean retweets and favorites not independent in rate.



In this figure I create bar chart between stages of doges and frequency of them.

You see that pupper have the most values, but the null data in this filed was alot, so it is not a big truth but give us insight about the most stage of doges in data.