We Rate Dogs on Twitter

Programm: Udacity Data Analyst Nanodegree (DAND)
Student: Laman Mammadli
Project: Data Wrangling

Introduction

After wrangling the data, I saved it as twitter_archive_master.csv file. Then read it by means of pandas library and loaded the data into twitter_master table. After making a few cleanings, I started to analyze it.

Data Analysis

Before starting the data analysis process, I write down the questions which I am interested to find answers. Below you can see the questions, analysis and observation:

Questions

1. What is highest and lowest retweet count tweet?
2. What is highest and lowest favorite count tweet?
3. What is the highest and lowest rating tweet?
4. Do favorite tweets retweet more?
5. Do favorite tweets and retweets have a relationship with the ratings of the tweets?

Analysis

1. I used 'idmin' and 'idmax' to find the index of lowest and highest tweet and created a function called find_minmax. And I used it to answer the other two questions as well.

Tweet which has the highest retweet_count :
Here's a doggo realizing you can stand in a pool. 13/10 enlightened af (vid by Tina Conrad) https://t.co/7wE9LTEXC4
Tweet which has the lowest retweet_count :
Oh my. Here you are seeing an Adobe Setter giving birth to twins!!! The world is an amazing place. 11/10 https://t.co/11LvqN4WLq

2. I used the function above to find the answer to my second question.

Tweet which has the highest favorite_count :
Here's a super supportive puppo participating in the Toronto #WomensMarch today. 13/10 https://t.co/nTz3FtorBc
Tweet which has the lowest favorite_count :
RT @dog_rates: This is Lilly. She just parallel barked. Kindly requests a reward now. 13/10 would pet so well https://t.co/SATN4If5H5
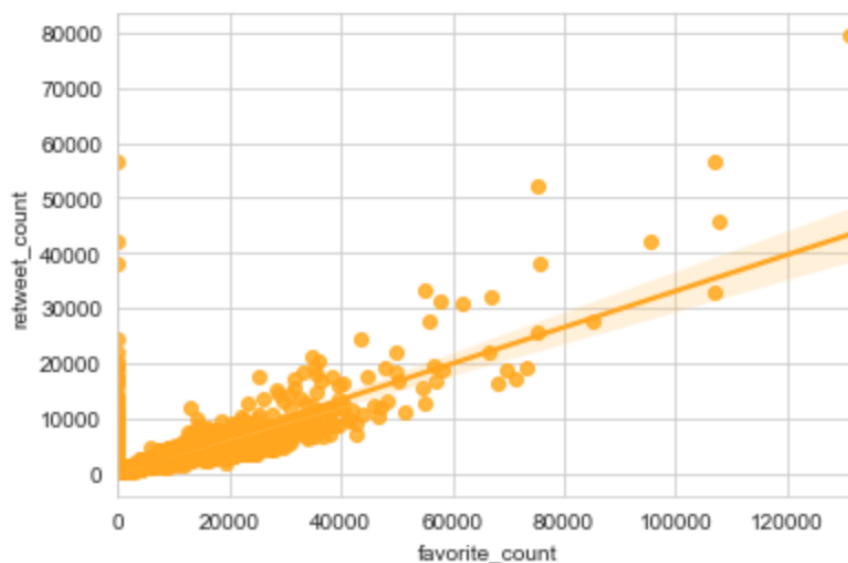
3. I used find_minmax() function to answer to my third question.

Tweet which has the highest rating_numerator :
This is Atticus. He's quite simply America af. 1776/10 https://t.co/GRXwMxLBkh
Tweet which has the lowest rating_numerator :   When you're so blinded by your systematic plagiarism that you forget what day it is. 0/10 https://t.co/YbEJPkg4Ag

4. I was interested in to find relation between favorite tweet and retweets. As both are the quantitative variables, I made a scatter plot. As seen from the plot below, there is a positive relationship between favorite tweets and retweets. It means that people retweet the favorite tweets.

5. In the last question, I was interested to check whether there is a relationship between ratings and retweets and favorite tweets. The plots display that there is not any relationship.