

# Data Wrangling (Twitter Account WeRateDogs)

## Analyzing and Visualizing report

In the other report ( wrangle report ) the twitter account dataset called WeRateDogs was gathered, cleaned, stored and assessed but in this report an analysis and visualization will be conducted on the dataset.

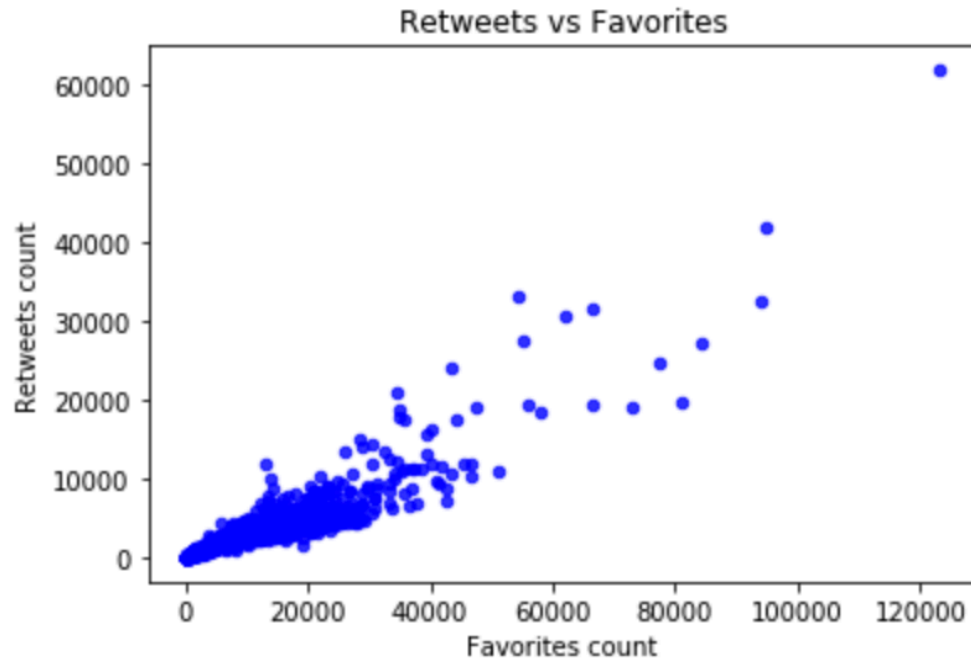
### Analyzing Data

First the merged csv file with all of the three datasets was read and saved in a dfMergeFinal variable then an analyzing was conducted through display the rows and description of the dataset also drop the tweet id because it is duplicated, display rating\_denominator with 10 and 1 and display retweet\_count greater than 100 or less than 100.

### Visualizing Data

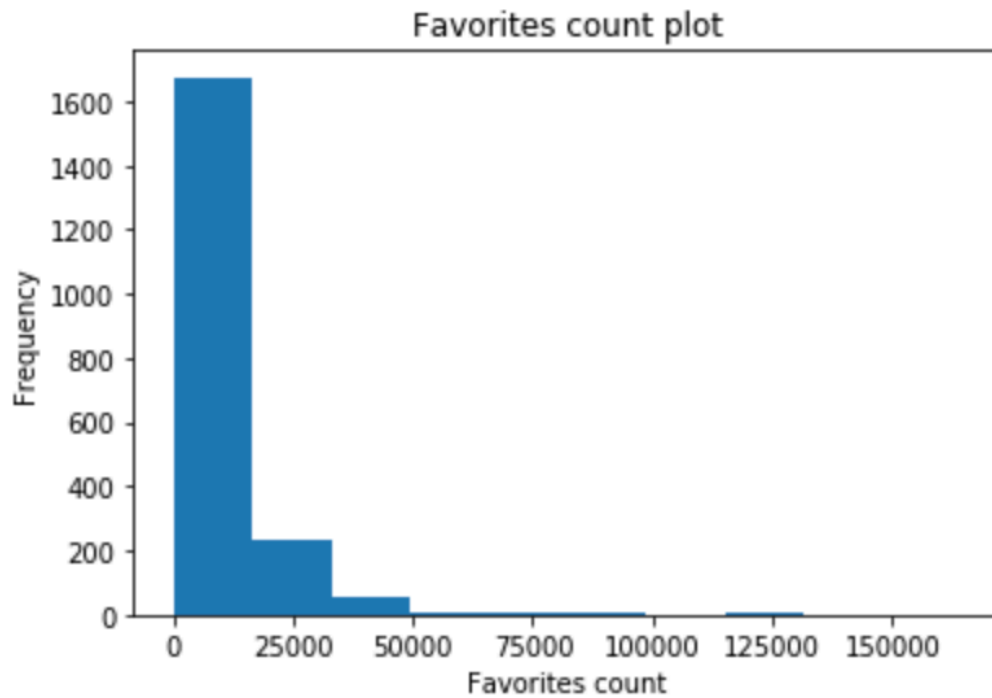
There are many plots to be conducted but two plots were chosen:

- 1) plot the favorites count vs Retweets count: that shows the number of retweets vs the number of favorites (likes) in this popular WeRateDogs account.



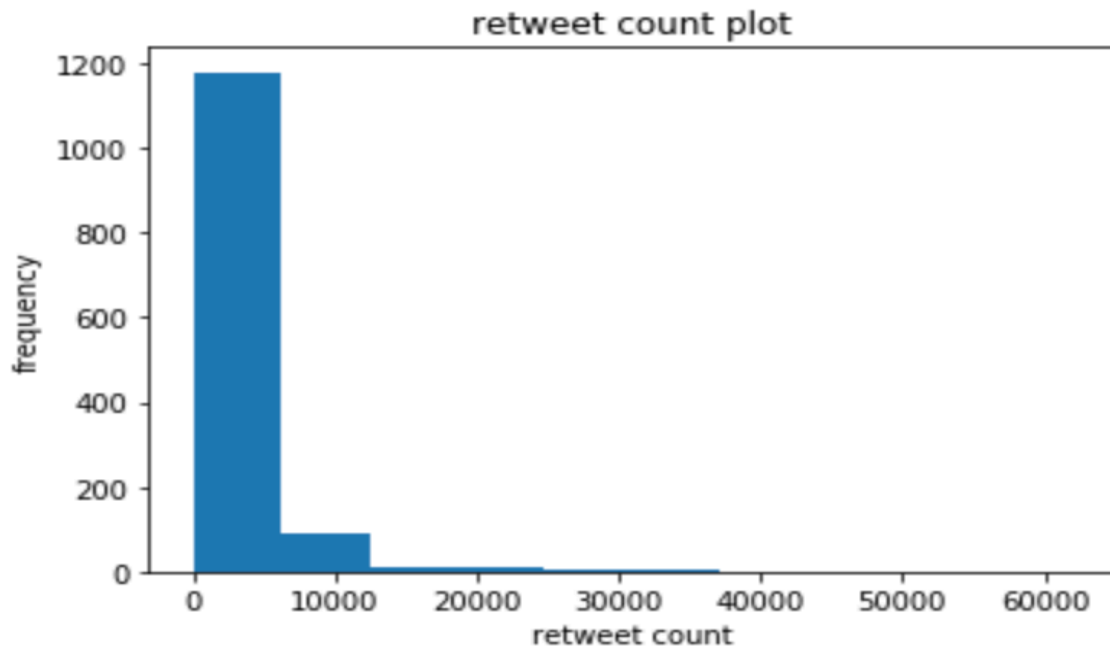
The above plot shows that when a tweet has many likes it has too many retweets because most people when they like they retweet too and this way it leads to increase in both retweets count and favorites count also it shows a strong positive correlation.

- 2) plot the favorite count: that shows the number of favorites (likes) in this popular WeRateDogs account.



This plot shows the favorites count (likes) on WeRateDogs Account range and it shows that the most popular favorites count is between ( 0 - 2000 ) and there is a little between ( 4000 - 120000 ) there for it is a descending favorites count.

3) plot the retweet count: that shows the number of retweets in this popular WeRateDogs account.



This plot shows the retweets count on WeRateDogs Account range and it shows that the most popular retweets count is between ( 0 - 1000 ) and after that it descend.