

Analyzing and Visualising Data

For the completion of

Udacity project two

By

Rashidat Sikiru

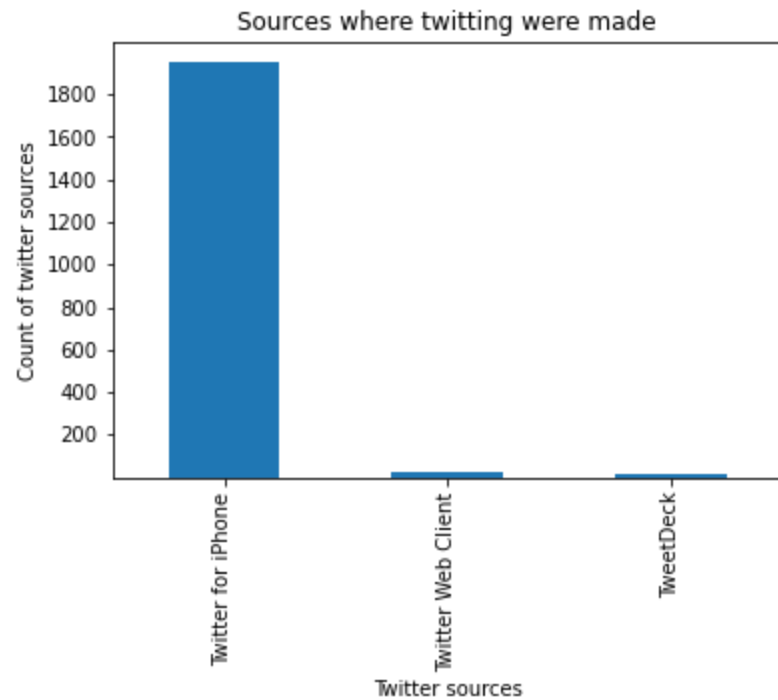
Introduction:

After the datasets were cleaned and merged together, it was then converted to a CSV file. To make it available for visualization, the CSV file was read in and then all columns whose data types changed were transformed back to its correct data type. To further the analysis, some questions were asked:

1. What is the most used twitter source?
2. What is the most common Dog name?
3. What is the relationship between favorite count and retweet count?

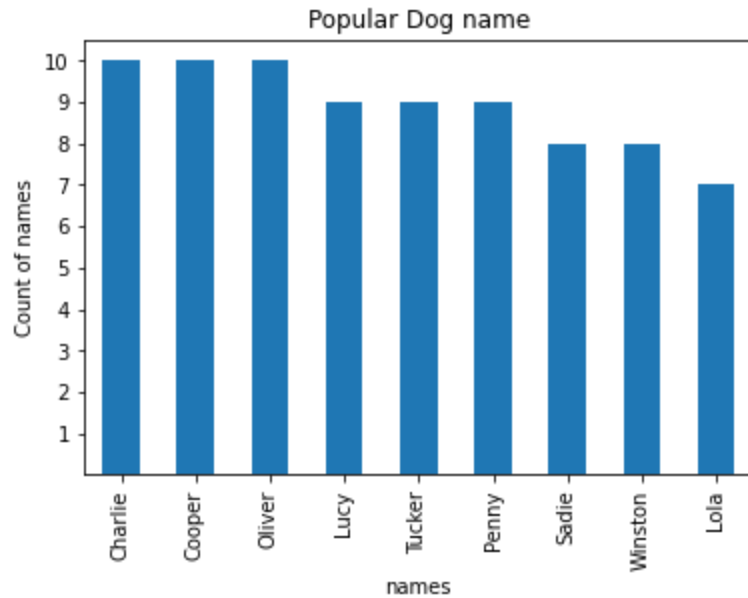
Q1. What is the most used twitter source?

The most common twitter source used is the twitter for iphone(~1900), followed by twitter web client and tweet deck. As shown in the figure below:



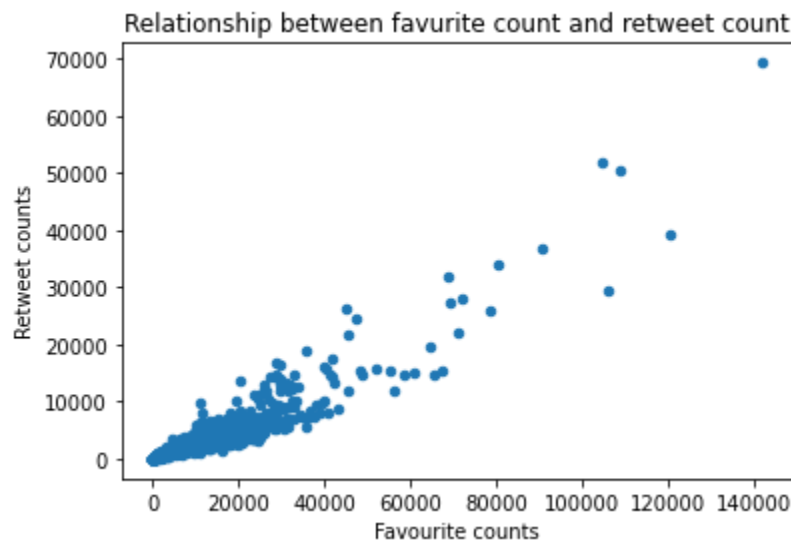
Q2. What is the most common Dog name?

The common names used are Charlie, Cooper and Oliver. Followed by a couple of others. This can be shown below



Q3. what is the relationship between favorite count and retweet count?

The analysis shows a positive correlation between favorite counts and retweet count. That is, the higher the number of favorites, the higher the number of retweet. This is represented below



Conclusion

The above report clearly shows the analysis carried out. It showed that the most popular tweet source is the twitter for iphone. It also clearly explained that the most popularly tweeted dog name is charlie, followed by cooper and others. Lastly, it also clearly illustrated a positive relationship between favorite counts and retweeted counts. Which explained that, the more the favorites, the more it is going to be retweeted. This is a limited research, further analysis can be carried out.