

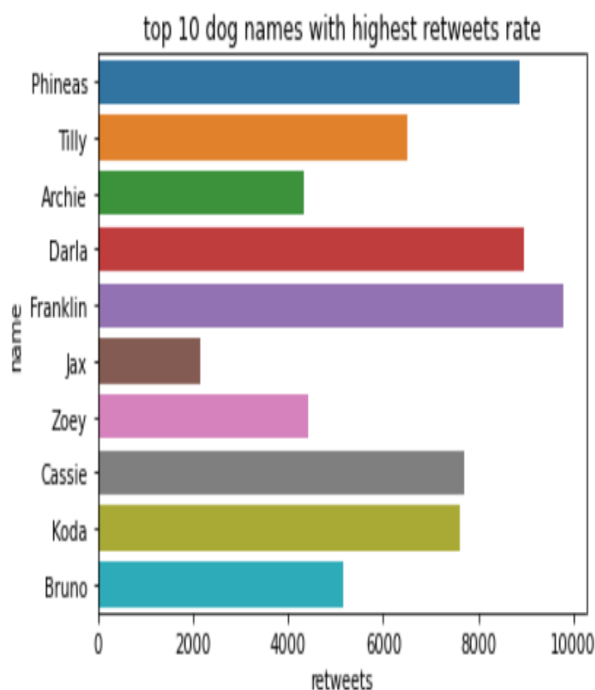
Project5: Wrangle and Analyze Data

Visualization And Insight:

We want to use visualizing to answer certain questions once we finish the wrangling process, which includes gathering, assessing, and cleaning the provided data. After gathering, assessing, and cleaning our data, we used several attributes such as :

- Number of retweets in Int
- Number of favorites in Int
- The name of dog after removing the unclear name
- The sources after removing unwanted data from it

Visualization 1: Find out the top ten dog names with retweets rate.

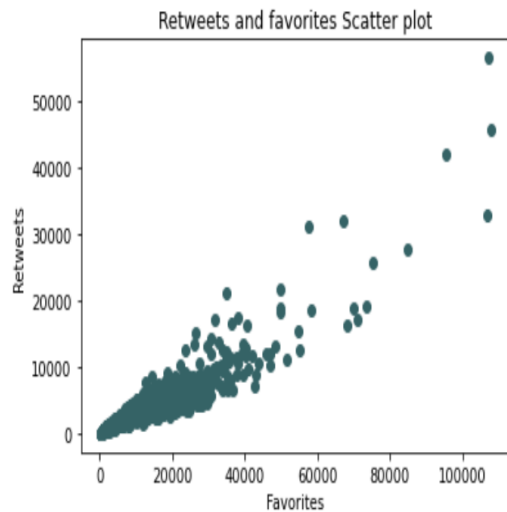


After removing None values and unsuitable dog names such as (the, very, all, a, an, all, the and so on.), also convert the Number of retweets to integer we aspire to find out the top 10 dog names with the greatest retweets rate value however, As seen in the graph, the name of the dog "Franklin" received the most retweets, followed by the name of the dog "Phineas." followed by the name of the dog "Darla"



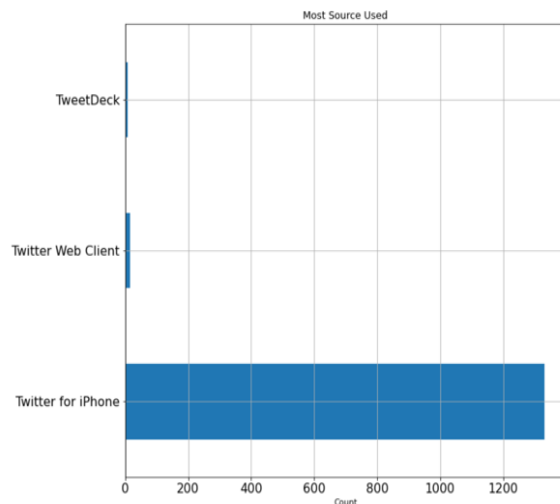
UDACITY

Visualization 2: Scatter plot between favorites and retweets.



We want to discover if there is a relation between favorites and retweets. As expected in the scatter graph there is a strong positive correlation (+0.92) shown by the correlation result between favorites and retweets. Therefore there is a strong positive correlation between them which indicates when the number of favorite goes up the retweets goes up to

Visualization 3: Most sources used



We can observe from the figure that Most sources used by iPhone, furthermore this is expected because handling Twitter application in the iPhone is easier than the web



UDACITY

sources

- <http://ocrpsychology2015.blogspot.com/2015/06/>
- https://video.udacity-data.com/topher/2018/November/5be5fb4c_twitter-api/twitter-api.py
- <https://towardsdatascience.com/twitter-analytics-weratedogs-a441be7d4a85>
- <https://www.earthdatascience.org/courses/use-data-open-source-python/intro-to-apis/twitter-data-in-python/>