act_report

August 29, 2022

0.1 Report: act_report

This dataset was a very interesting one and a lot of analysis could be made based on almost every column. However, three (3) questions were addressed and analysis were made. These questions were

1. The source of most tweets and ratings
2. The relationship between favorite counts and retweet counts
3. The highest dog rating

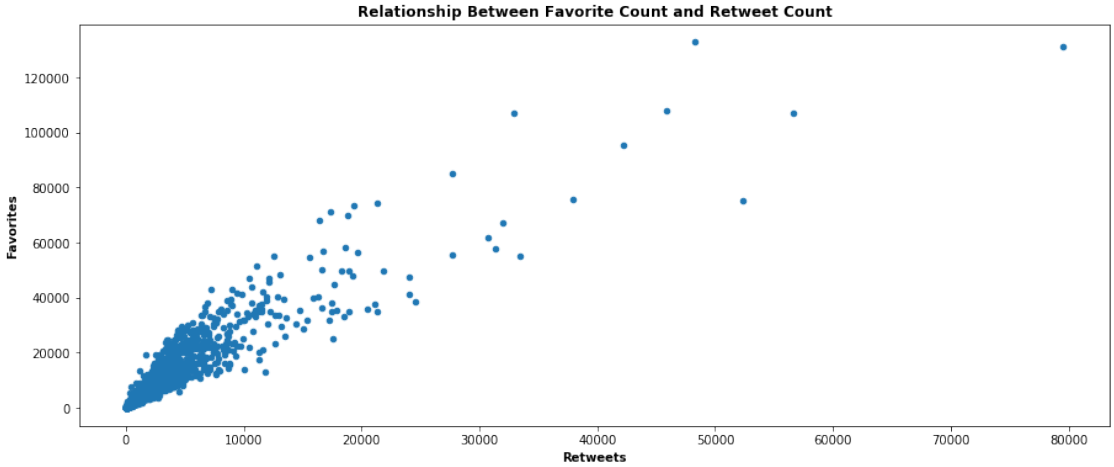
For the first question; the source of most tweets and ratings, the dataset was viewed and the source value counts was assigned to a variable which was 'rating_source'. The assinged variable(rating_source) was then run to determine the source with the highest value count. The various sources were shown; Twitter for iphone, Vine - Make a scene, Twitter Web Client and TweetDeck. The highest source was Twitter for iphone with 1961 counts.

The relationship between favorite and retweet counts was of interest in the second analysis. Matplotlib and seaborn were imported for the plotting of a scatter plot. The ratings were then plotted with a figsize of 15, 6 with 'Relationship between favorite count and retweet count as the title.The xlabel was Retweets and the ylabel indicated the Favorites. After plotting, scatter plot showed a strong correlation between the ratings. The conclusion draw was tweets with the most likes are likely to be retweeted more. The scatter plot is shown below

For the final analysis, we were concerned about the highest rating. Which numerator rating had the highest number? Since the denominator was fixed as 10, the numerator range was of interest and so was the particular rating with the highest count. The value counts for the various numerator ratings were checked. The ratings ranged from 0 to 16, 17, 26 and 27. The 12 numerator rating had the highest value count of 487

Visualization of the relationship between favorite count and retweet count

In [3]:



Tweets with the most likes are likely to be retweeted more.