act_report

August 2, 2022

0.1 Report: act_report

- Create a **250-word-minimum written report** called "act_report.pdf" or "act_report.html" that communicates the insights and displays the visualization(s) produced from your wrangled data. This is to be framed as an external document, like a blog post or magazine article, for example.

1 Report on analysis and visualization of wrangled data

The data used for this process was obtained from three sets of data. These data include

- A comma seperated file : 'twitter-archive-enhanced.csv'
- A tab seperated file : 'image-predictions.tsv' and
- A text file : 'tweet-json.txt'

I cleaned and merged the data to a master dataframe that enabled me to assess and gain insights into the data for analysis. The data is about different dog ratings as well as other information including images, comments on the dog and the name of the dog gotten from twitter.

1.0.1 Insights and Analysis

From the assessment of the master dataframe, I was able to pose some questions that the data could provide. These include

- What tweets did the users like the most and their ratings?
  - This question came from viewing the statistics of the favorite_count in the dataframe. I noticed that the highest favorited tweet had a count of 132810 and if it had anything to do with ratings. From my analysis, I noticed that amongst the top 20 favorited tweets, the most liked or favorited tweets were of dogs with at least a 10/10 rating with the most frequent being 13/10 and the highest being 14/10.

- What source are most tweets gotten from?
  - This came from the extraction of the categories from the source column during the cleaning phase of the wrangling process. I noticed that there were four categories namely: Twitter for iPhone, Vine, Twitter Web Client, TweetDeck with Twitter for iPhone having the most count of 2042 and TweetDeck having 11 counts. From this we can conclude that most users tweet from the iPhone.