

WRANGLE REPORT

Wrangle and Analyze Data Project

In this project Data were collected from three different sources.

1. I started with the twitter-archive-enhanced.csv Data Frame.
 - I gathered it then assessed it then clean it.
2. The second Data Frame I needed Twitter API to gather it by using python tweepy library, so I had to make twitter developer account to the API that I need and I extracted the favorite and retweet counts for each tweet. This data was then saved as a JSON file.
 - Then I load it in jupyter notebook as Data Frame
 - The Assess and Clean.
3. The third Data Frame was image_predictions.
I needed requests.get(url) to download it from the internet, the I load it in jupyter notebook.
 - And I did the same steps Assess then Clean

the reason why I did each three steps (Gathering – Assessing – Cleaning) separately for every Data Frame so it's will be easy to understand the whole process and I made a copy for each Data Frame before start cleaning.

I used several functions and methods already exists in python libraries to assess and clean the data and other I had to search for them in internet and edit the function to meet my purpose of using