Branch: **master** ▾   **udacity-data-wrangling-project** /
wrangle_project_final.ipynb

Find file   Copy path

S-Tabor Uploading Jupyter Notebook   6a03196   on Oct 26, 2018

1 contributor

3760 lines (3759 sloc)   124 KB   ‹›   ▤   Raw   Blame   History   🖥   ✏   🗑

Data Wrangling Project: WeRateDogs Twitter Data

```
In [703]: import pandas as pd
          import numpy as np
          import requests
          import tweepy
          import json
          import matplotlib.pyplot as plt
```

Gathering Data

1) Twitter archive data:

```
In [704]: twitter_archive = pd.read_csv('twitter_archive_enhanced.c
          sv', sep=',')
```

2) Image prediction data:

```
In [10]: # Downloading and saving data using Requests
         url = 'https://d17h27t6h515a5.cloudfront.net/topher/2017/A
         ugust/599fd2ad_image-predictions/image-predictions.tsv'
         r = requests.get(url)
         open('image_predictions.tsv', 'wb').write(r.content)
```

```
Out[10]: 335079
```

```
In [705]: image_predictions = pd.read_csv('image_predictions.tsv',
          sep = '\t')
```

3) Retweet/Favourite count data

```
In [7]: # Downloading data using Twitter API

        consumer_key = 'HIDDEN'
        consumer_secret = 'HIDDEN'
        access_token = 'HIDDEN'
        access_secret = 'HIDDEN'

        auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
        auth.set_access_token(access_token, access_secret)

        api = tweepy.API(auth, wait_on_rate_limit = True, wait_on_
        rate_limit_notify = True)
```