

MINI PROJECT

I have retrieved some data from twitter using twitter API for a query like #mh17 and created a bar chart between countries (x-axis) and number of tweets for query = "#mh17"(Y-axis).

Language Used: Python

Two programs:

1. twitter_analysis.py
2. read_from_csv.py

Steps which I have followed:

- Created an account and app on <https://apps.twitter.com/> and got consumer keys and access tokens.
consumer_key = 'XXXXXXXXXX'
consumer_secret = 'XXXXXXXXXX'
access_token = 'XXXXXXXXXXXXXXXXXXXX'
access_token_secret = 'XXXXXXXXXXXXXXXXXXXX'

1. **twitter_analysis.py** (run : python twitter_analysis.py)
 - a. Python package which need to installed to use these programs : tweepy, geopy, numpy, matplotlib and dependencies
 - b. Used OAuth Handler to authorized twitter app using the keys and tokens.
 - c. **api = tweepy.API (auth):** Creation of the actual interface, using authentication
 - d. To get all the attributes of the user of tweet used api.search() method with filtered parameters
 - e. Above search function returns JSON Object
 - f. Retrieved user attributes from JSON object of each tweet (like timestamp, status, location)
 - g. It returns location from user profile if users share the location to their tweets
 - h. From location I have got longitude and latitude using geopy module
 - i. From longitude and latitude I have got country name
 - j. All attributes stored in **result.csv** file.
2. **read_from_csv.py** (run : python read_from_csv.py)
 - a. Opened and read **result.csv** file using csv module
 - b. From .csv file I have count the frequencies of each countries.

- c. And stored country name as key and frequency of that country as value in python dictionary.
- d. And made a plot between them using matplotlib python module.

3. Screenshot of Plots:

