

# Twitter Scraper

## 1 Introduction

In this assignment, I scraped tweets using the Tweepy API, and search queries were based on recent trending topics. Here I collected 5000 tweets on the trending topic #BlackHistoryMonth and did an organic performance analysis.

## 2 Data Statistics

I reported Twitter's organic performance matrices i.e Retweet Count, Favorite Count, Followers Count, Friends Count, and geotags for every user. As an example, here I am reporting some interesting statistics for user\_id [24123226]:

- Retweet Count = 0
- Favorite Count = 2
- Followers Count = 4961
- Friends Count = 5455
- Location = Houston TX

## 3 Successes

1. Collected 5000 tweets and saved them in the data\_full.json file after removing retweets and tweets with less than 5 words.
2. Randomly collected 300 samples. These data were then saved in the data\_sample.json file.
3. Reported twitter's organic performance metrics such as retweet count, favorite count, followers count, friends count, and geotags.
4. Script details have been described in the readme.txt file.

## References

As a reference, I used the following tutorial: Use Tweepy & Python to Scrape the Twitter API