

Pothi.com Programming Task

Use the Twitter Streaming API to track a given keyword and generate various reports about the tweets.

Input

keyword (program should prompt on the command line)

Task 1

After starting, every 1 minute, program should print following reports based on the tweets received in last 1 minute.

User Report

Print the names of all Twitter users who tweeted with given keyword along with the count of tweets from them.

Links Report

Print the total number of links included in the tweets followed by a list of the unique domains sorted by the count. Any shortened links should be expanded.

Content Report

Print the number of unique words used in the tweets followed by the list of top 10 words sorted by occurrence. Remove the common function words like a, an, the, of, with etc. You don't need to use a NLP package. Filtering based on a list of function words like articles, prepositions is sufficient.

Task 2

Print the same reports every 1 minute as above but now they should be based on last 5 minutes of data instead of last 1 min.

When the program starts, the report printed at 1 min should be based on data in the first minute, report at 2 min should be based on data of first 2 minutes and so on. Reports after 5 minutes should be based on the data from last 5 minutes. For example, the report printed on the completion of 11th min, should be based on the data received between 6th-11th minutes.

Notes

- You can complete the task in programming language of your choice.

Submission

Setup a local Git repo with your code. The repo should include a README file with instructions on how to setup

and run the code. Make a compressed .zip or .tar.gz archive of the repo and send it as email attachment.

Resources

- [Twitter Streaming API Docs](#)