ASSIGNMENT 2

For this assignment, you will use Twitter Search API to download tweets about anything that you care about. For example, you could search for "UTD" and get a list of tweets. Additionally, you have to search for the same topic for 6 different timelines to get 6 different files. For example, you can search for UTD on 6 different days to get 6 input files for MapReduce program to process. The MapReduce program will then look for Hashtags i.e. those words starting with "#" to find which Hashtags are trending for a topic.

Here are the requirements:

1. You have to use the API i.e. sign up for a developer account and use a programming language (preferably Java, but not required) to search for tweets. You might find some useful information here:

https://dev.twitter.com/rest/public/search

2. You have to search for the same topic on 6 different timelines. It is up to you to choose time intervals e.g. one day, one week, etc. But you need to get 6 medium sized input files containing tweets. Remember to search for a topic and not a hashtag.

3. After getting the 6 input files, you can run the MapReduce code on them. *The only change that you need to make is that instead of emitting a key value pair for each word, you have to look for Hashtags i.e. words starting with "#" and emit key value pair for them. You only need to make this change in the map method of the mapper class.*

4. Get the output of reducer as usual.

What to submit:

1. Your code to get tweets. Do not turn in your input files as this will consume a lot of space on eLearning.

2. Your modified MapReduce code.

3. The output of MapReduce code containing trending Hashtags.

4. A Readme file indicating what topic you searched for, which timelines you used, and what was the size of the input files.

All this should be turned in on eLearning before class on Friday February 5.