CS 495: Introduction to Web Science

Fall 2013

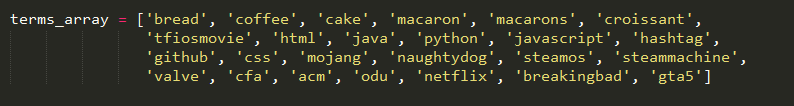
Assignment 2

Onapha Rattanachottiteepakorn

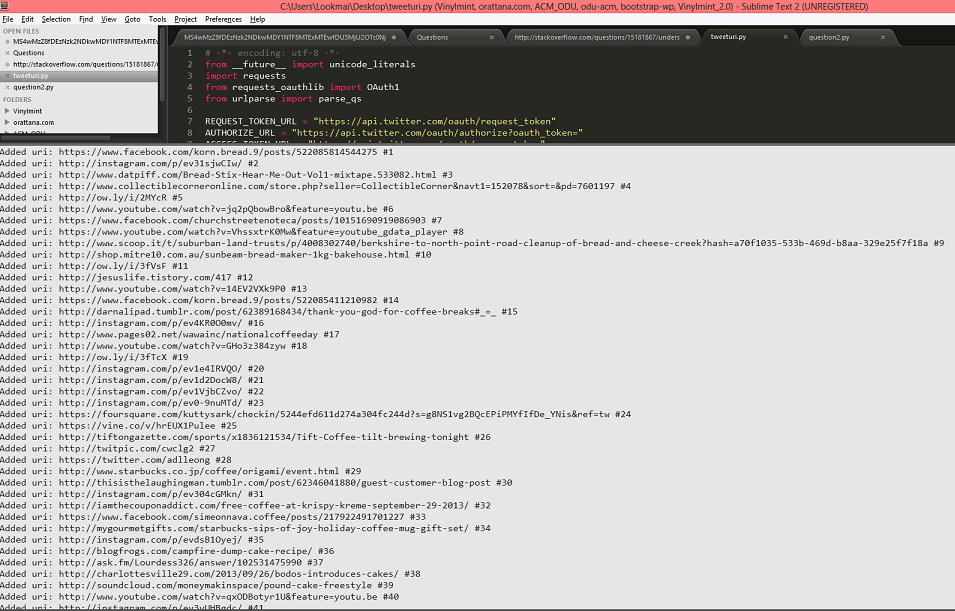
Submission Date: September 26, 2013

**Question 1**: Write a Python program that extracts 1000 unique links from Twitter. You might want to take a look at: <http://thomassileo.com/blog/2013/01/25/using-twitter-rest-api-v1-dot-1-with-python/>

The program requires user to have specific keys in order to get information from Twitter. To create a query, I used array of string that would use Twitter Search API to search the site with the keywords entered through the loop. The program returns search result and processes in order to determine if it is a URI and if it is a unique link. The search function repeats but the second time is to search through the hashtag query. After the search query is finished, the links are recorded into the file twitter\_links.txt.



**Figure 1**: array of keywords



**Figure 2**: example of the program running

**Question 2**: Download the TimeMaps for each of the target URIs. We'll use the ODU Memento Aggregator, so for example:

* URI-R = <http://www.cs.odu.edu/>
* URI-T = <http://mementoproxy.cs.odu.edu/aggr/timemap/link/http://www.cs.odu.edu/>

Create a histogram of URIs vs. number of Mementos (as computed from the TimeMaps). For example, 100 URIs with 0 Mementos, 300 URIs with 1 Memento, 400 URIs with 2 Mementos, etc.

The program processes through the text file twitter\_links.txt saved from assignment one. I couldn’t get the program to fully function, as the result of running through it returns 0 for every URI.

**Question 3**: Estimate the age of each of the 1000 URIs using the "Carbon Date" tool:

<http://ws-dl.blogspot.com/2013/04/2013-04-19-carbon-dating-web.html>

Note: you'll have to download the tool and install; don't try to use the web service.

For URIs that have > 0 Mementos and an estimated creation date, create a graph with age (in days) on one axis and number of mementos on the other.

DERODEROPDPPERPDPERPDFPERP

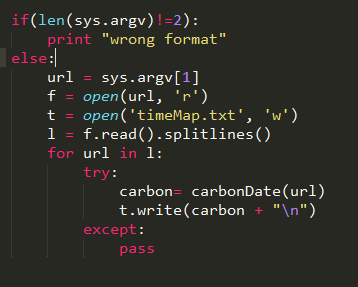


Figure 3: Sending in the URI to be carbondated.

**Sources**:

<http://stackoverflow.com/questions/15181867/understanding-the-set-function>

<https://dev.twitter.com/docs/using-search>

<https://dev.twitter.com/docs/api/1/get/search>

<http://thomassileo.com/blog/2013/01/25/using-twitter-rest-api-v1-dot-1-with-python/>

<http://stackoverflow.com/questions/4664850/find-all-occurrences-of-a-substring-in-python>