Data Mining Twitter

Abstract

Explains our project which is that finding current Georgia State University’s image using Twitter data.

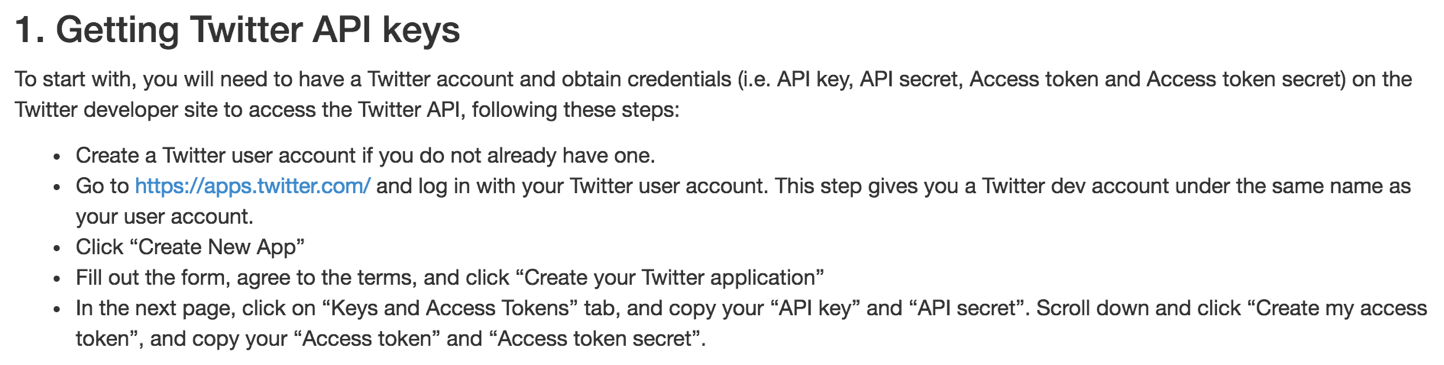
Twitter API

social-metrics.org/twitter-user-data/

socialmedia-class.org/twittertutorial.html

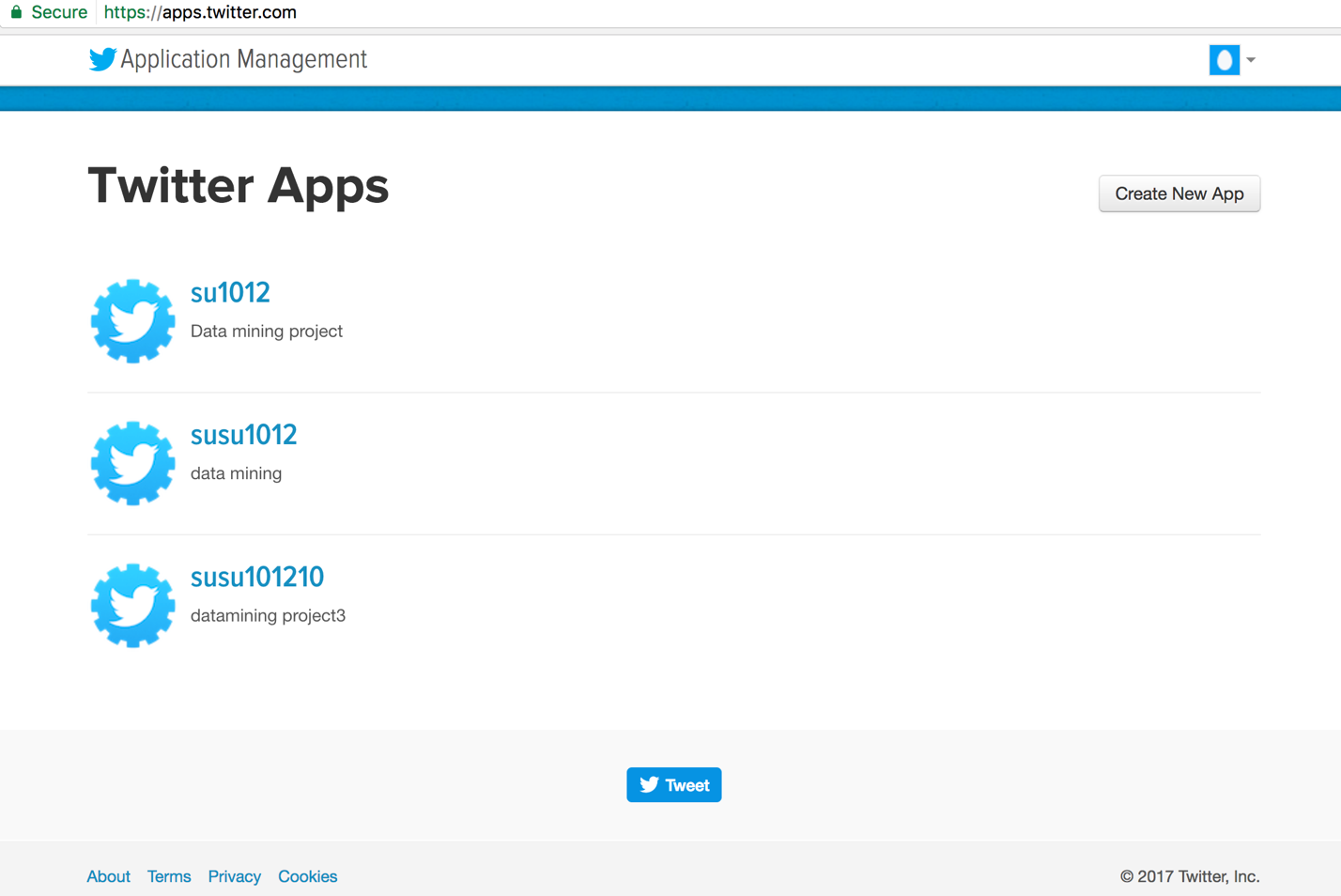
http://www-scf.usc.edu/~aupadhya/Mining.pdf

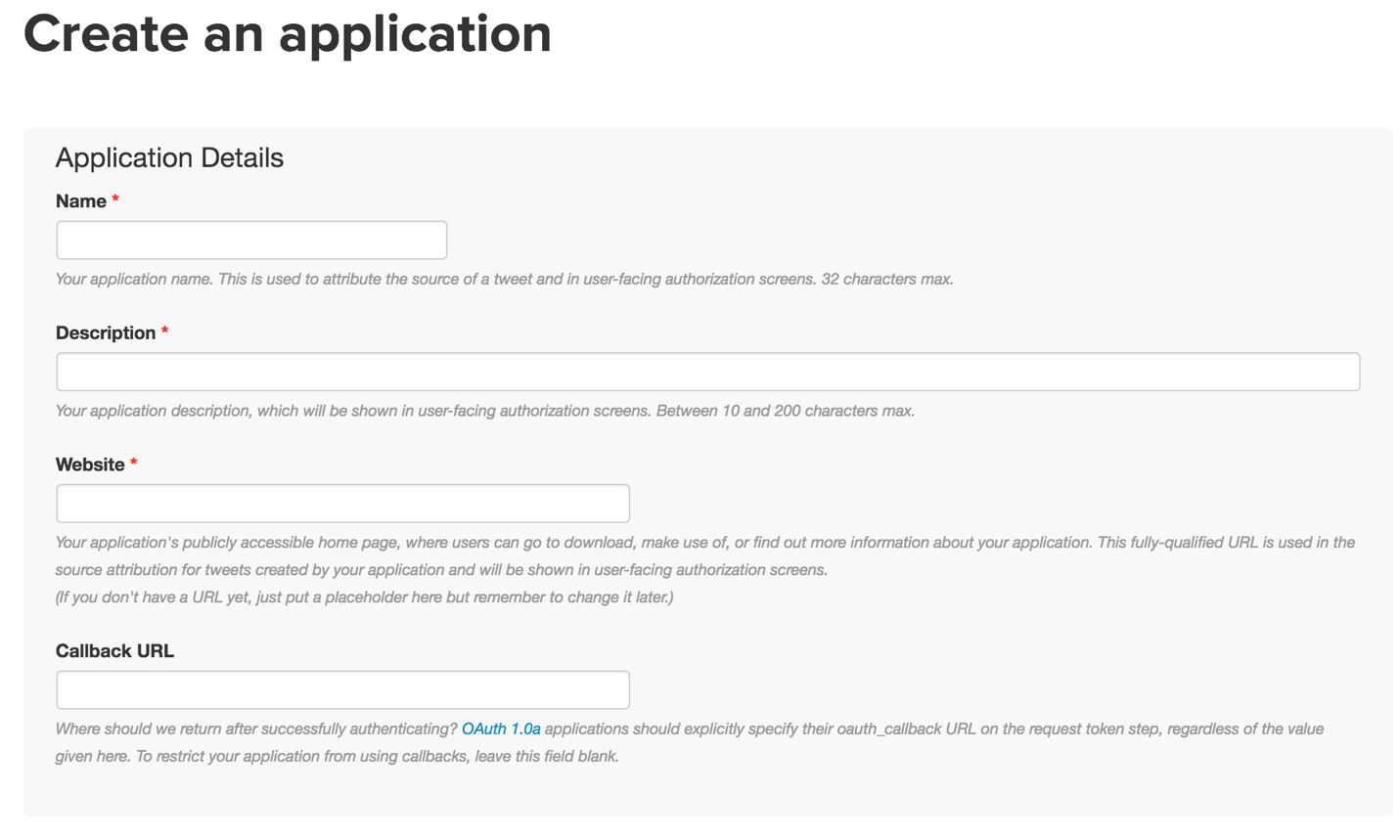
Please check above two links



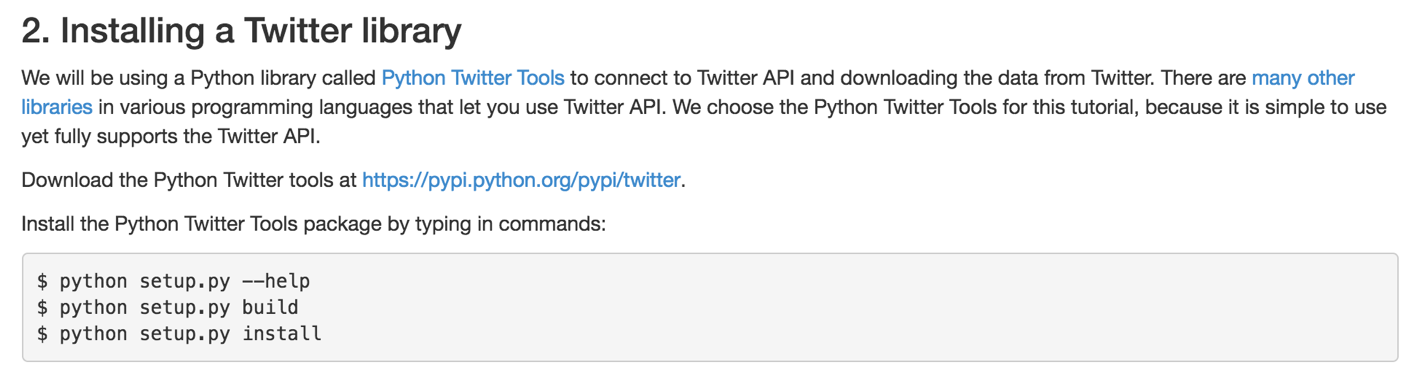
Above image is from the website

Access Token



This is my screen shot

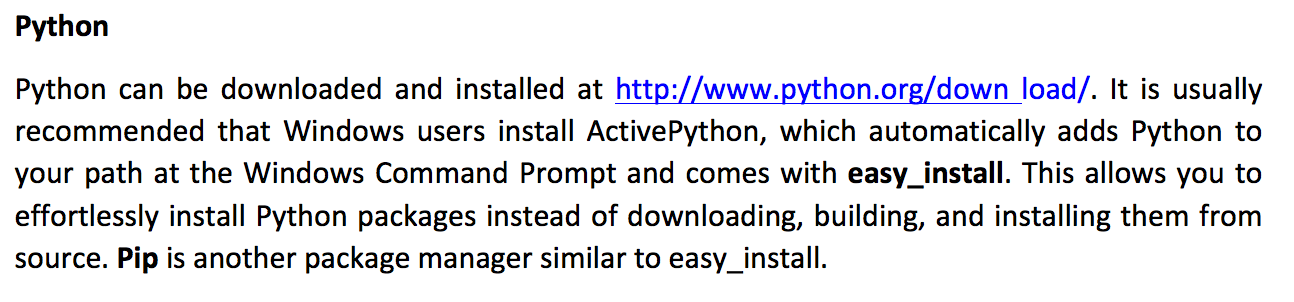
This is my screen shot



Above image is from the website

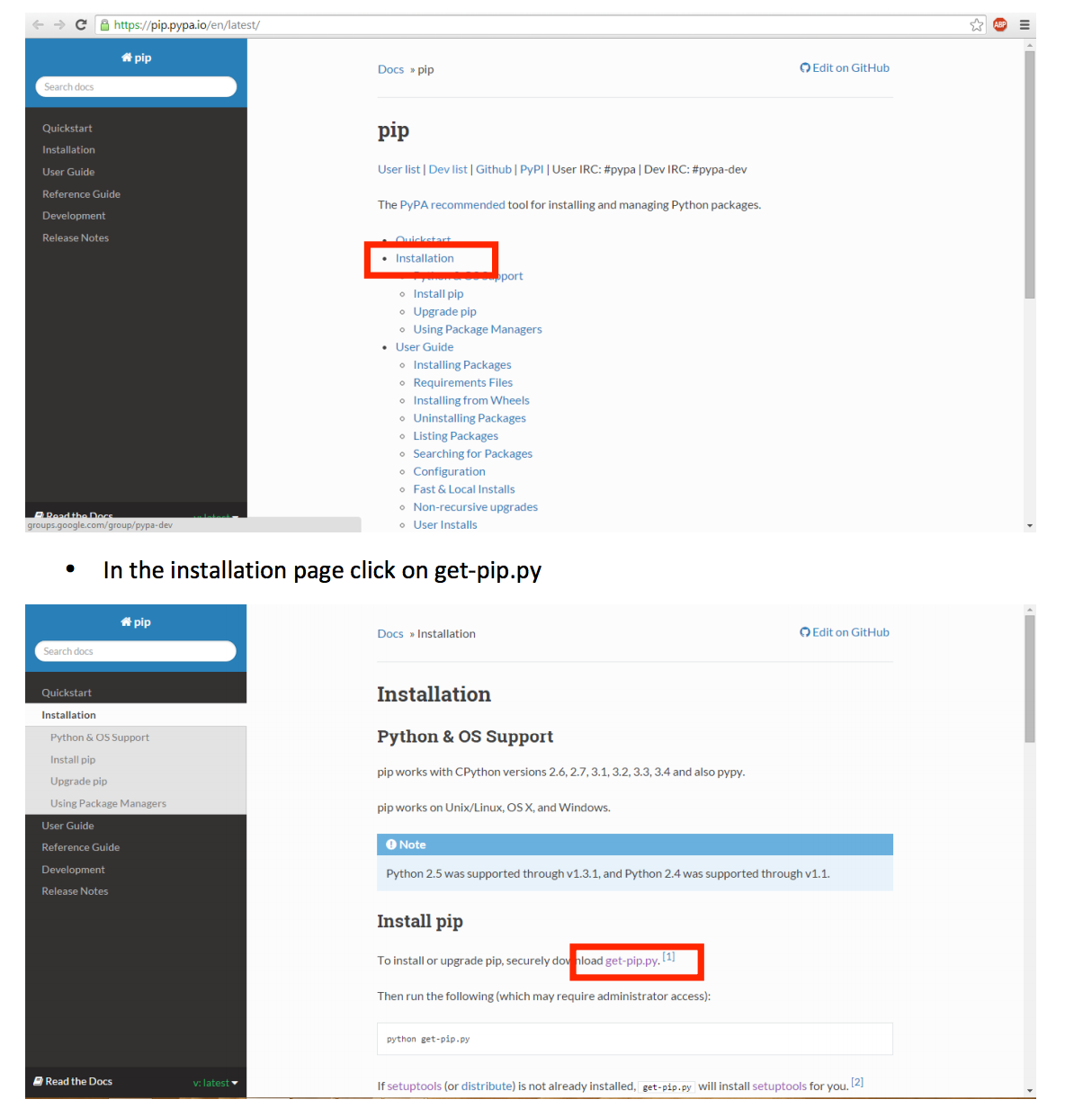
Tool Needed

1. Python

I setup python 3.6 which is the latest version.

Above image is from the website

<https://pypi.python.org/pypi/python-twitter/3.2.1> (I downloaded python-twitter 3.2.1 from this website) ( I think you can add a screen shot from this website)



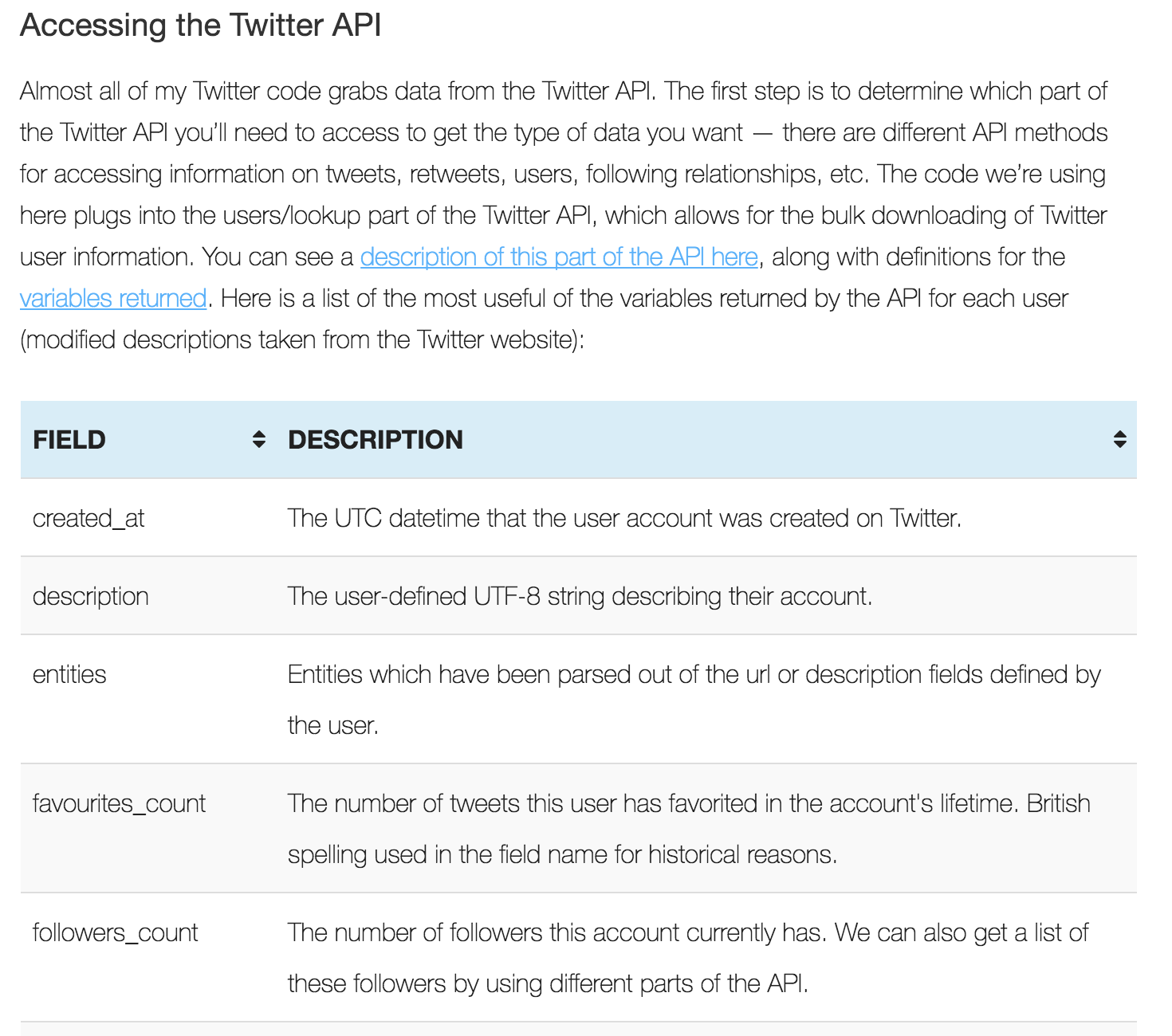
On the terminal

$ easy\_install pip

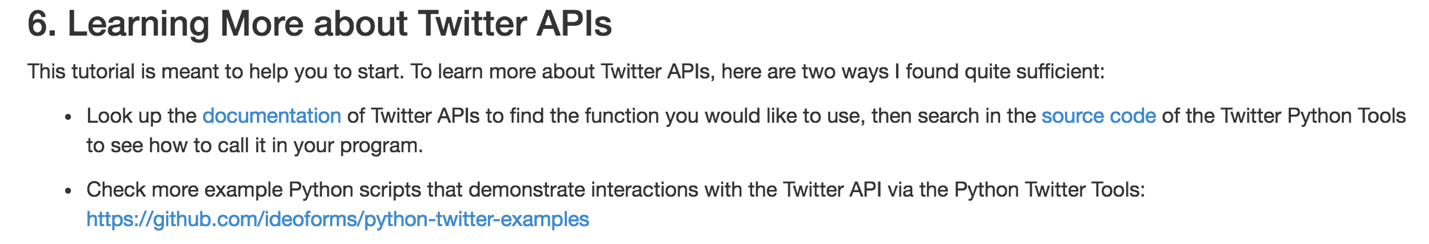
$ python get-pip.py

$ pip install simplejson

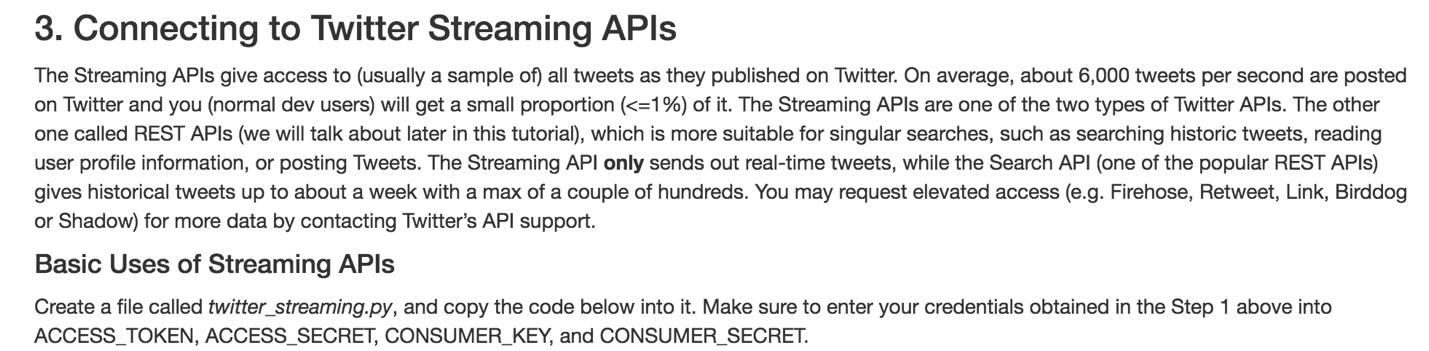
$ pip install twitter



Above image is from the website



Above image is from the website



Above image is from the website

I used Twitter Streaming API.

Code



Above code is to collect data by ‘walmart’.

On terminal

$ python collect.py > collect.json

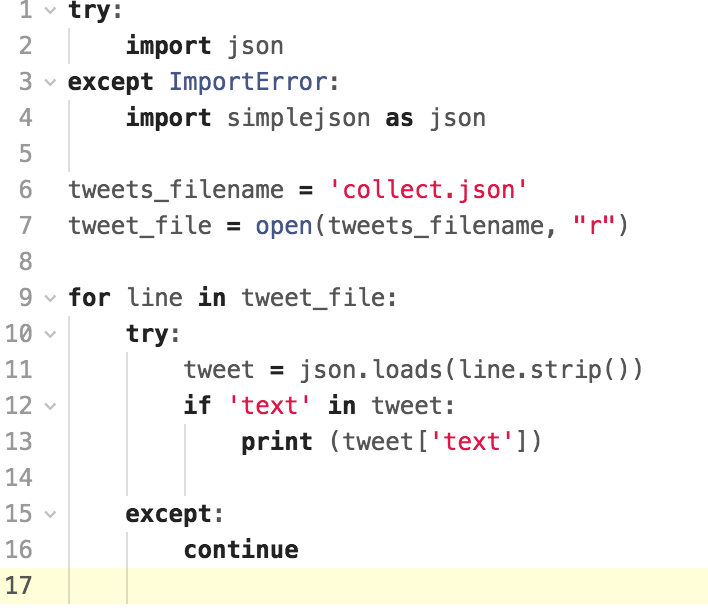
All data will save into collect.json

../../../Desktop/Screen%20Shot%202017-03-16%20at%201.13.41%20A

Above image is showing the file size.

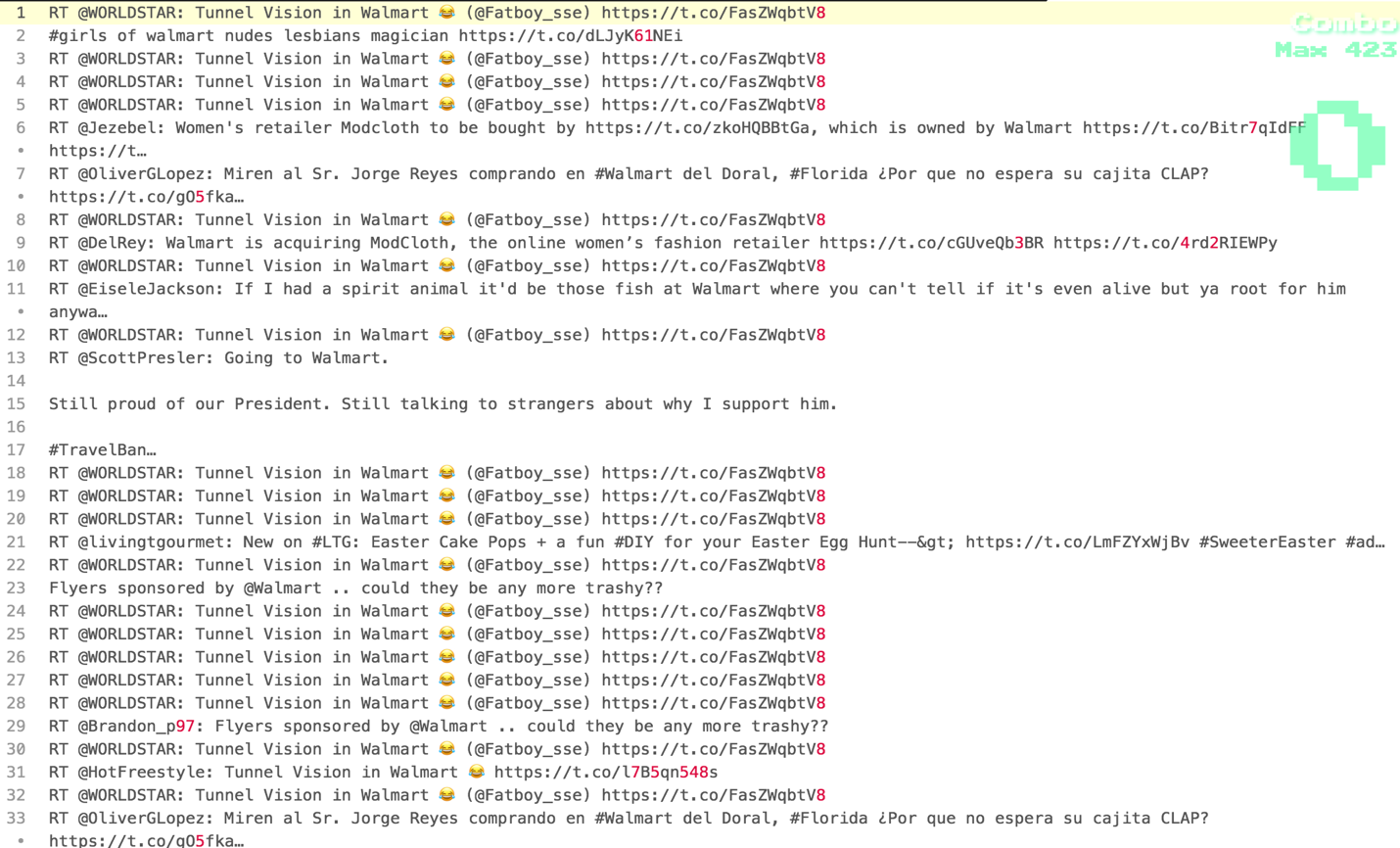


Above image is the result of collect.json.



Above code extracts only text.

$python extractText.py > extractText.json



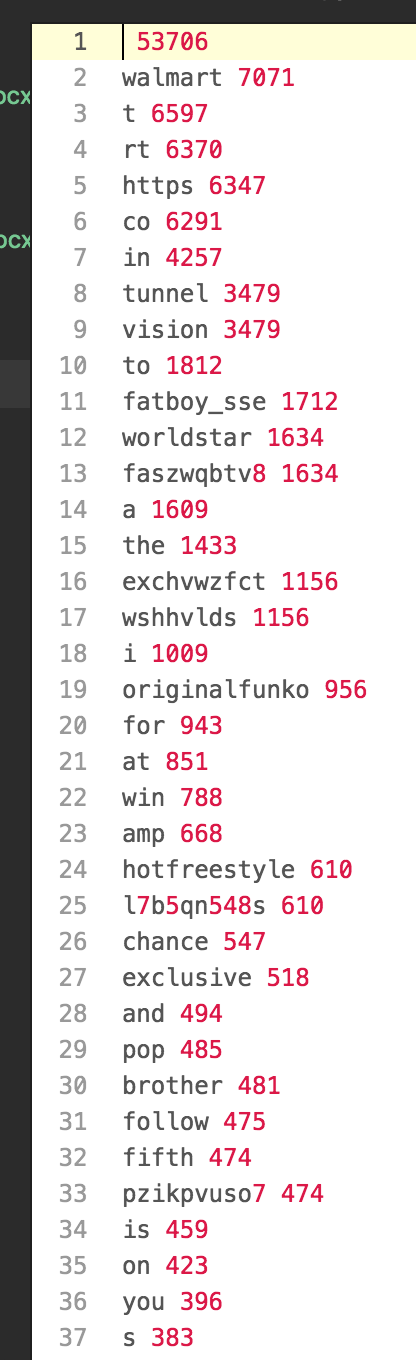
Above image is the result of extractText.json.



Above code is find out occurrence words top 500 words.

$ python ranking.py > ranking.json

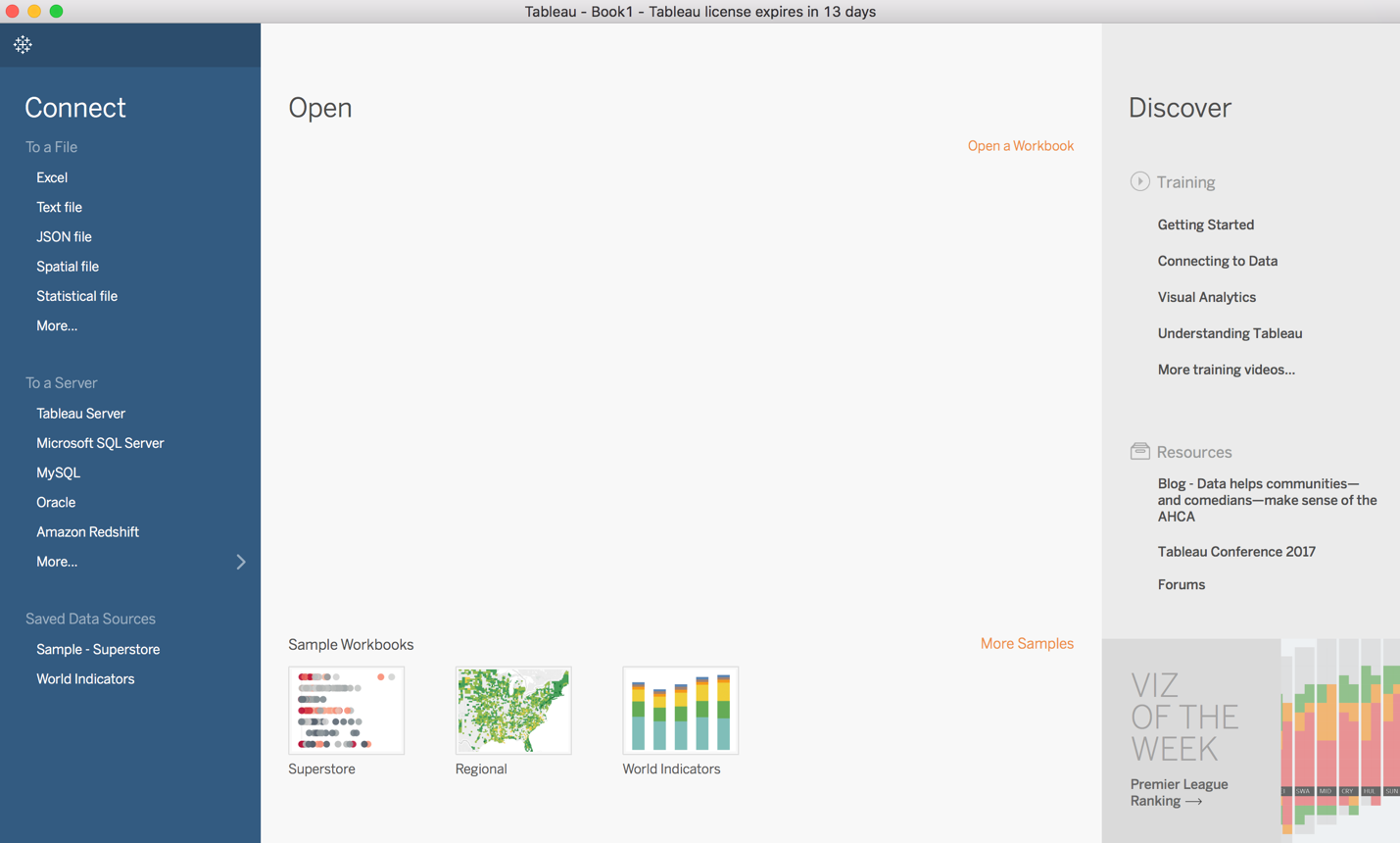
All result data will be stored into ranking.json.



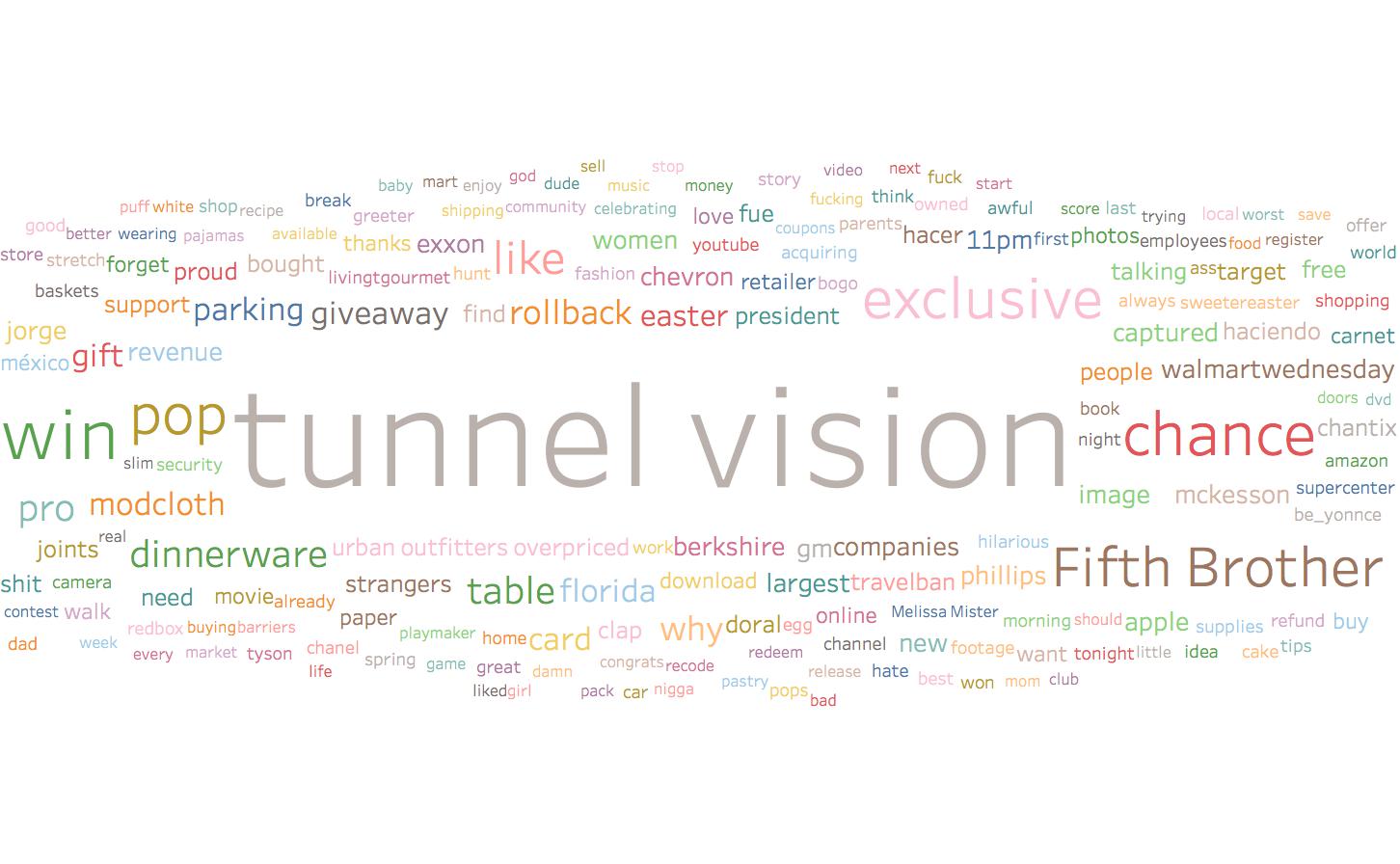
Above image is result of ranking.json.

I eliminated unimportant word like a, the etc. from ranking.json, then I stored in the excel file to make wordcloud image.

I installed tableau from <https://www.tableau.com/> to create wordcloud image file.



Above image is Tableau screenshot.



Above image is word cloud style of data visualization.

Conculsion