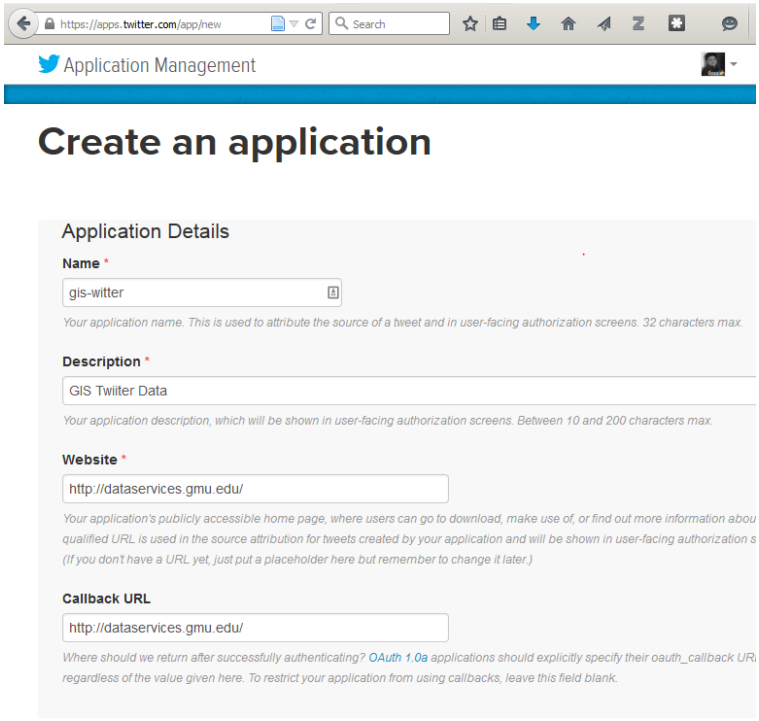
Getting Started with Twitter Data Collection

Note: For this exercises you will need some programming experience and should know how to install libraries. We will need **tweepy (http://www.tweepy.org/)** *library to install on your computer. You also need get API Keys in your twitter account prior to collect data from twitter. We have installed these on computers in the lab.*

The aim of this tutorial is to collect data from twitter with "GeoTagged" information by a keyword(s) and by a geographic area and visualize them using ArcGIS in CSV format.

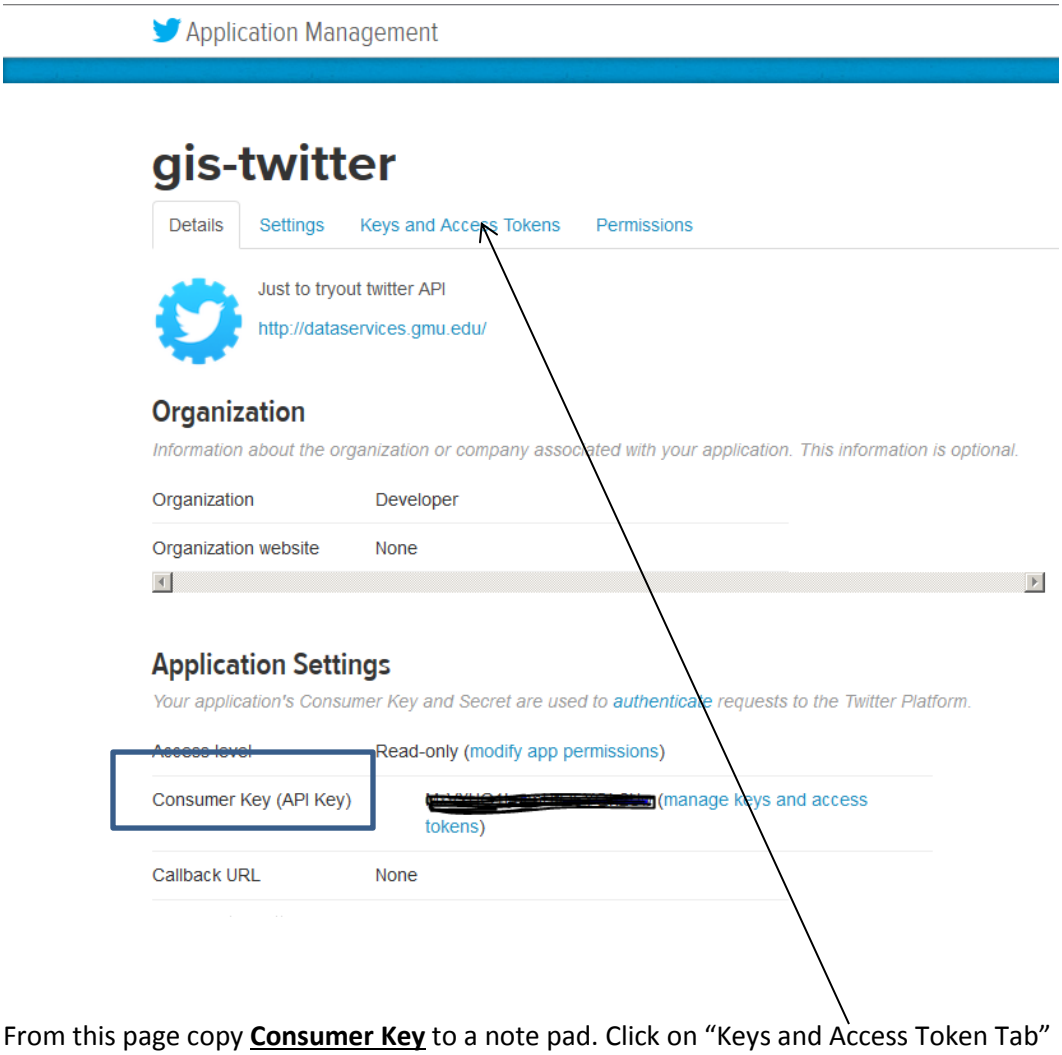
Steps

1. You need to create account at twitter.com. Create it now (at https://twitter.com/signup) if you don't have it now.
2. Once you have a working account, we need few tokens from twitter.
    a. Go to https://apps.twitter.com/app/new



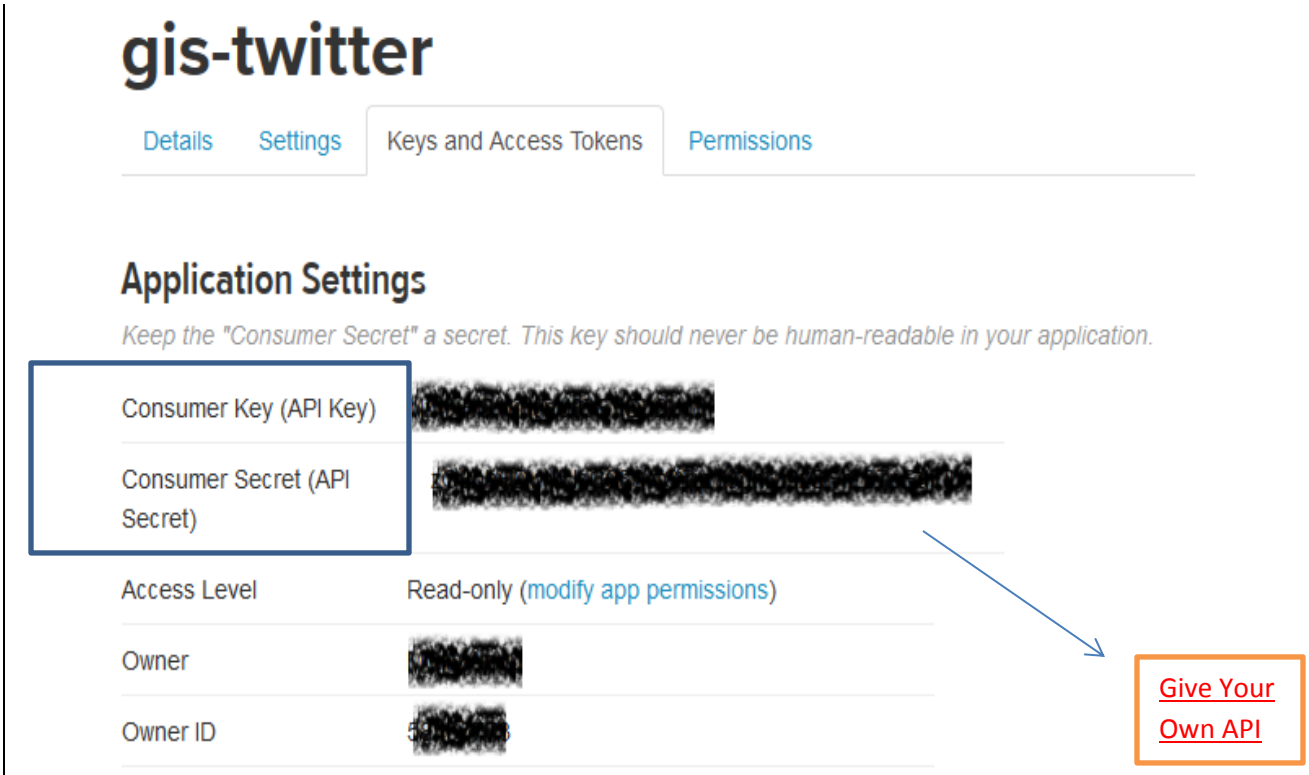
    b. Enter a name for the script that we are going to come up with, for this exercise we need not worry about website and callback url. Use any valid URL as shown above. Agree to the terms and click create application

3. After creation, application management page is displayed.

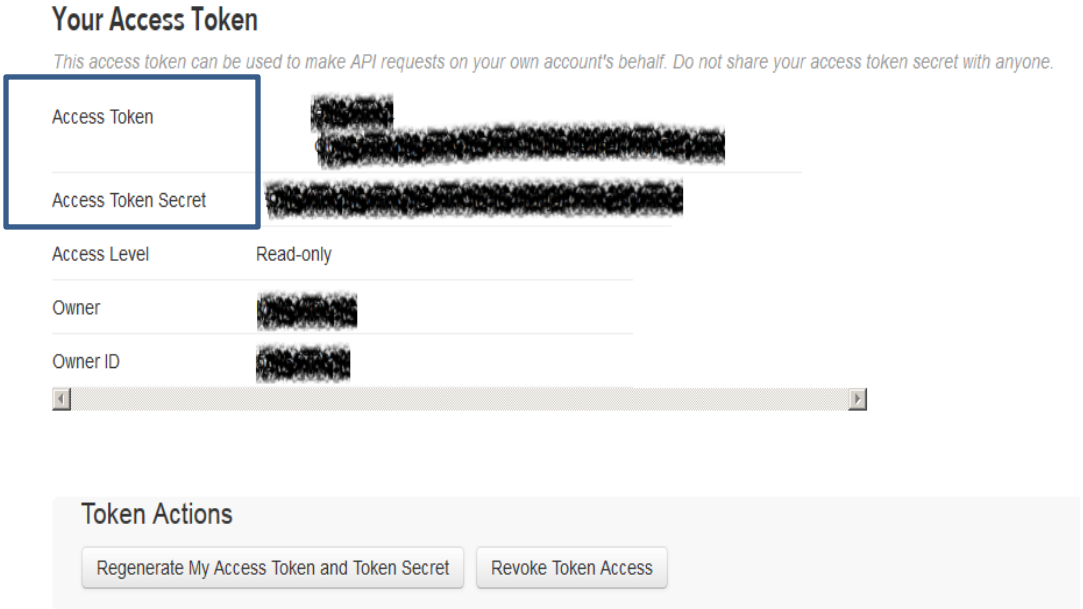


From this page copy **Consumer Key** to a note pad. Click on “Keys and Access Token Tab”

4. From Keys And Access Tokens, Copy **Consumer Key, Consumer Secret**

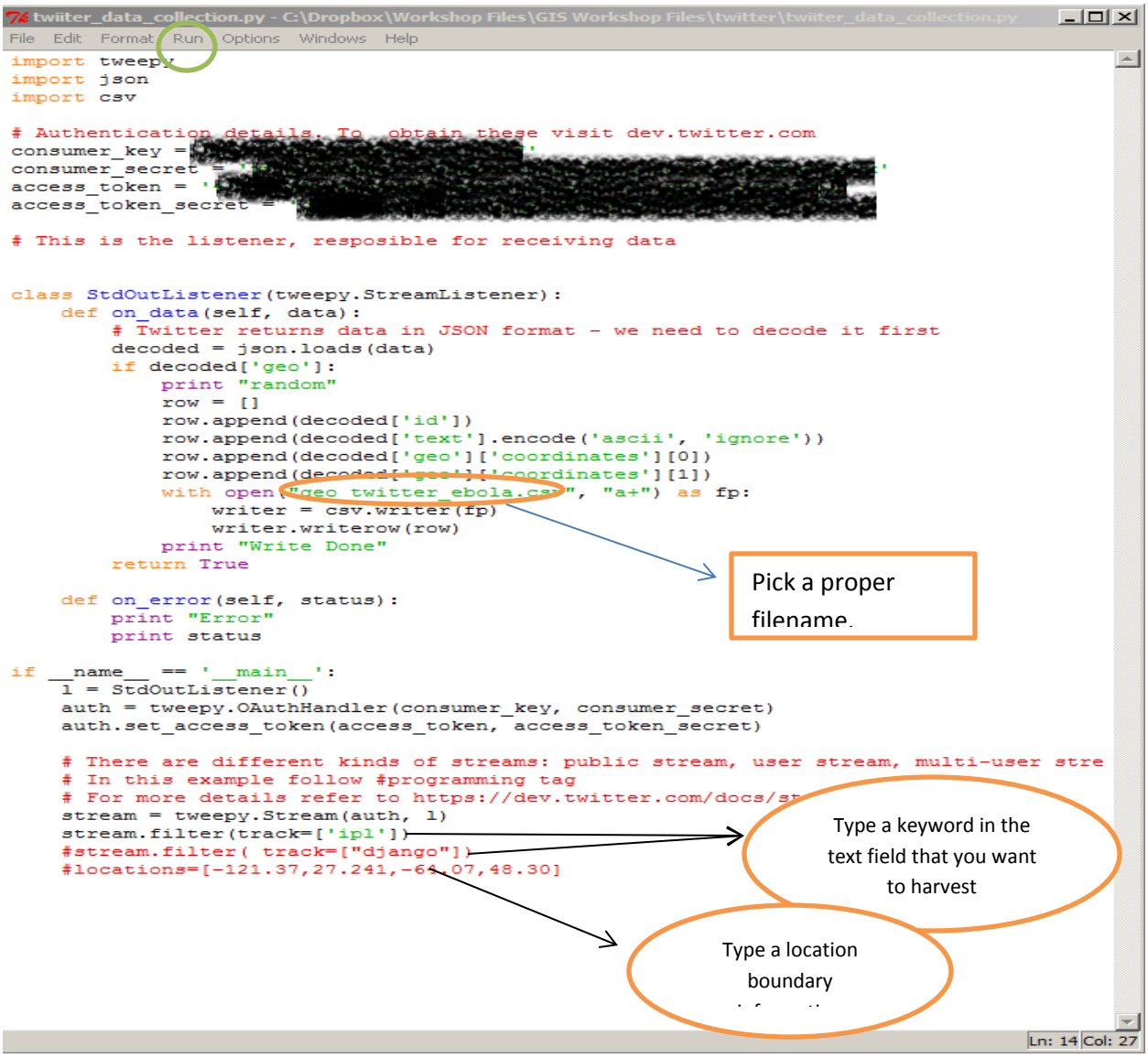


5. Now scroll down the same page, will have something as shown below> copy **Access Token**, **Access Token Secret**



6. Now we have everything that is needed to get started . Go ahead and download the script from https://gist.githubusercontent.com/anupkalburgi/227944dfc5b618e3d087/raw/be17a7685e7636ab8909cb0a3bb344e3a5998adf/twiiter_data_collection

   After downloading the file, save it to twitter_data_collection.py. And open up in "Idle" editor (in Python)and put in the keys that were copied from the twitter site



After getting the variables right, got to **Run** (green circle) and click on Run Module. That should start the stream. You might see SSL related warnings coming up, we don't worry about those for this given exercise

NOTE: This will bring you #hashtag word (eg, #django) and any word with "Django"

Locations field is to filter a geographic boundary box to populate twitter data.