**Twitter Scrapping Program – Instructions**

Note:- While opening python programs for making changes, use wordpad for better indentation.

**Prerequisites:-**

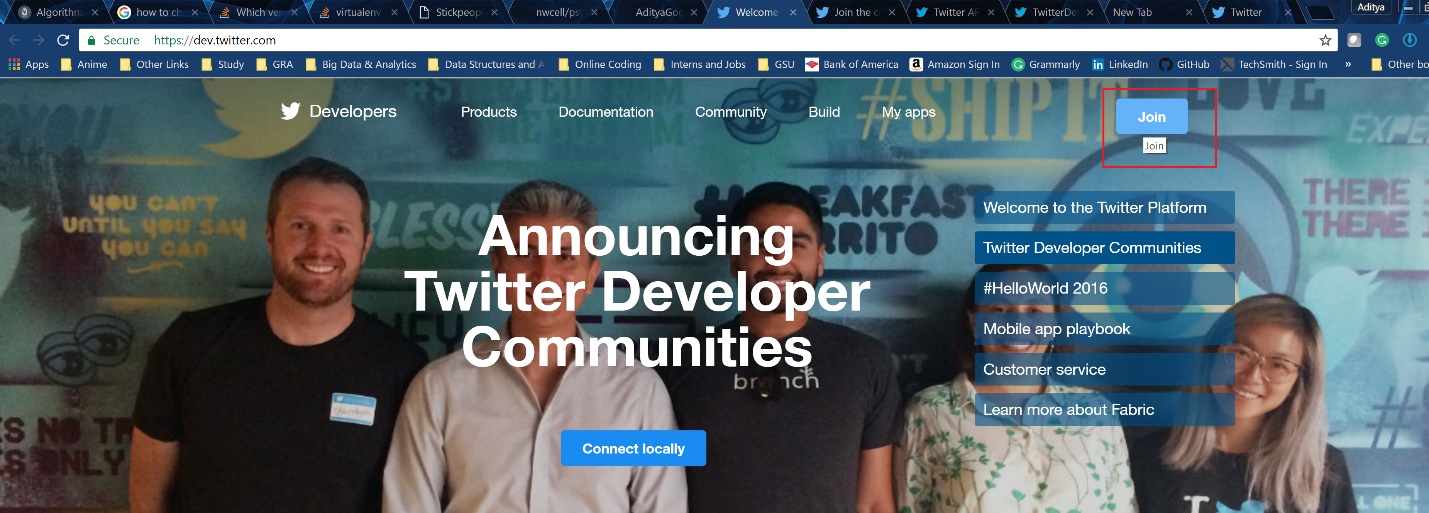
1. Create a Twitter account.

2. Sign in Twitter.

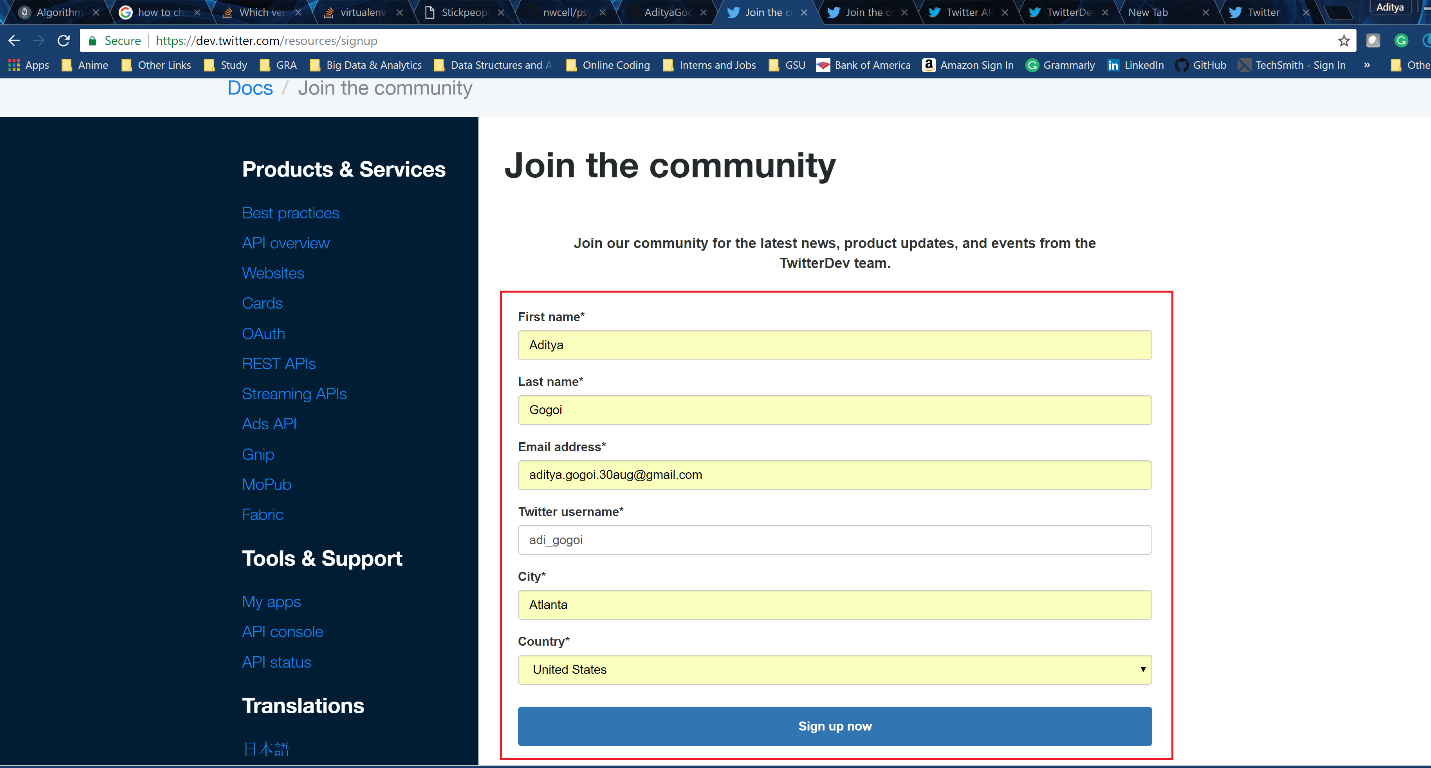
**Create a Twitter developer Account:-**

1. Go to the link :- <https://dev.twitter.com/>

2. Click the Join button.



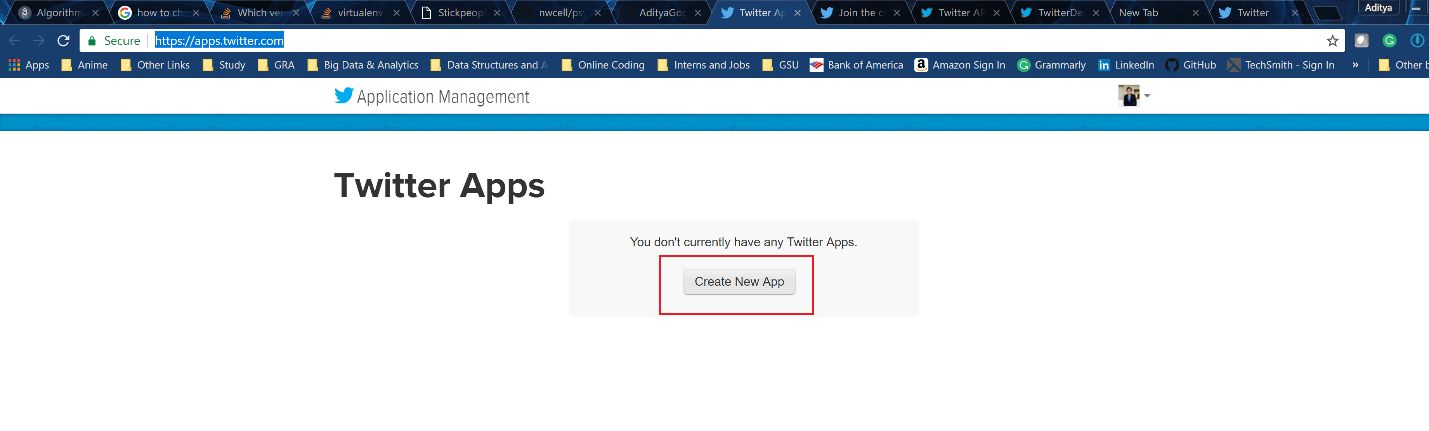
3. Sign up for the developer account with your details. After joining you will get a “Thank you for subscribing” message.



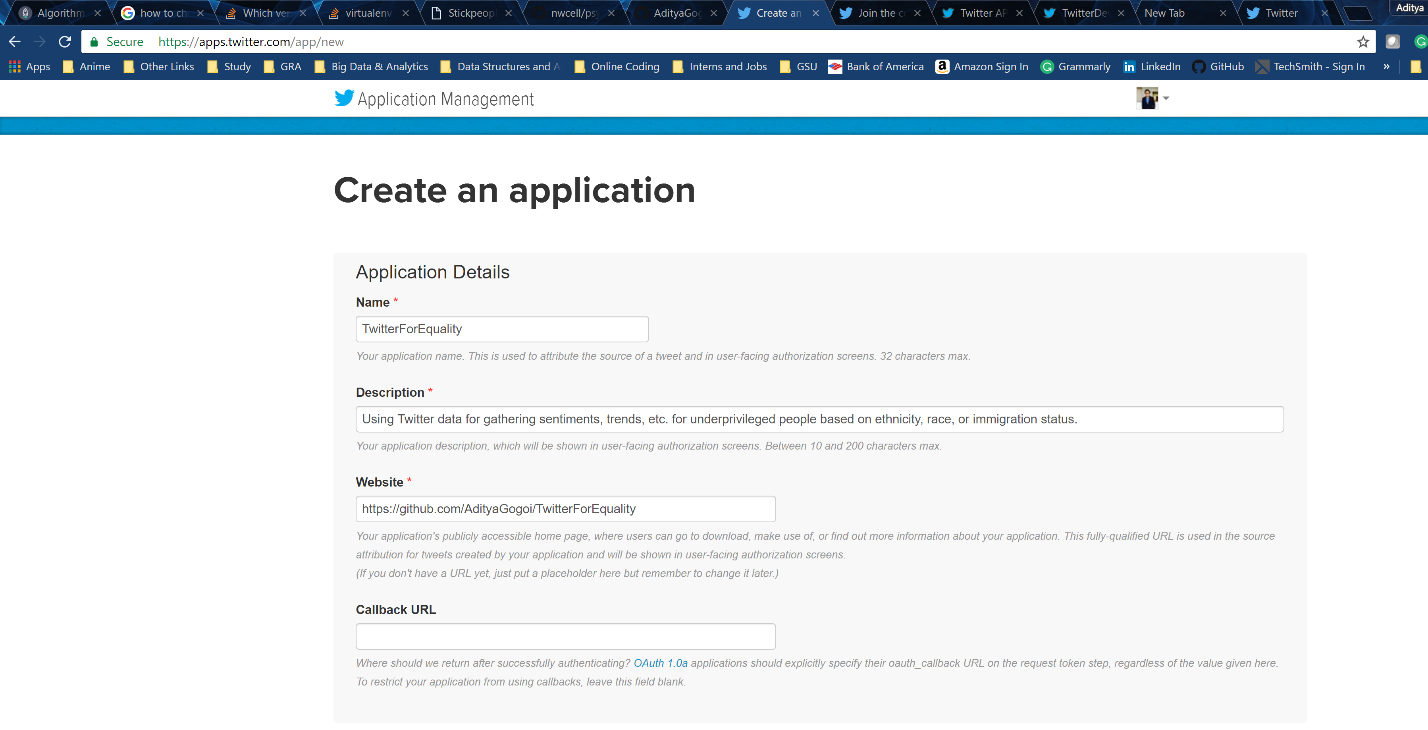
**Building your app on twitter:-**

1. Go to the link:- <https://apps.twitter.com/>

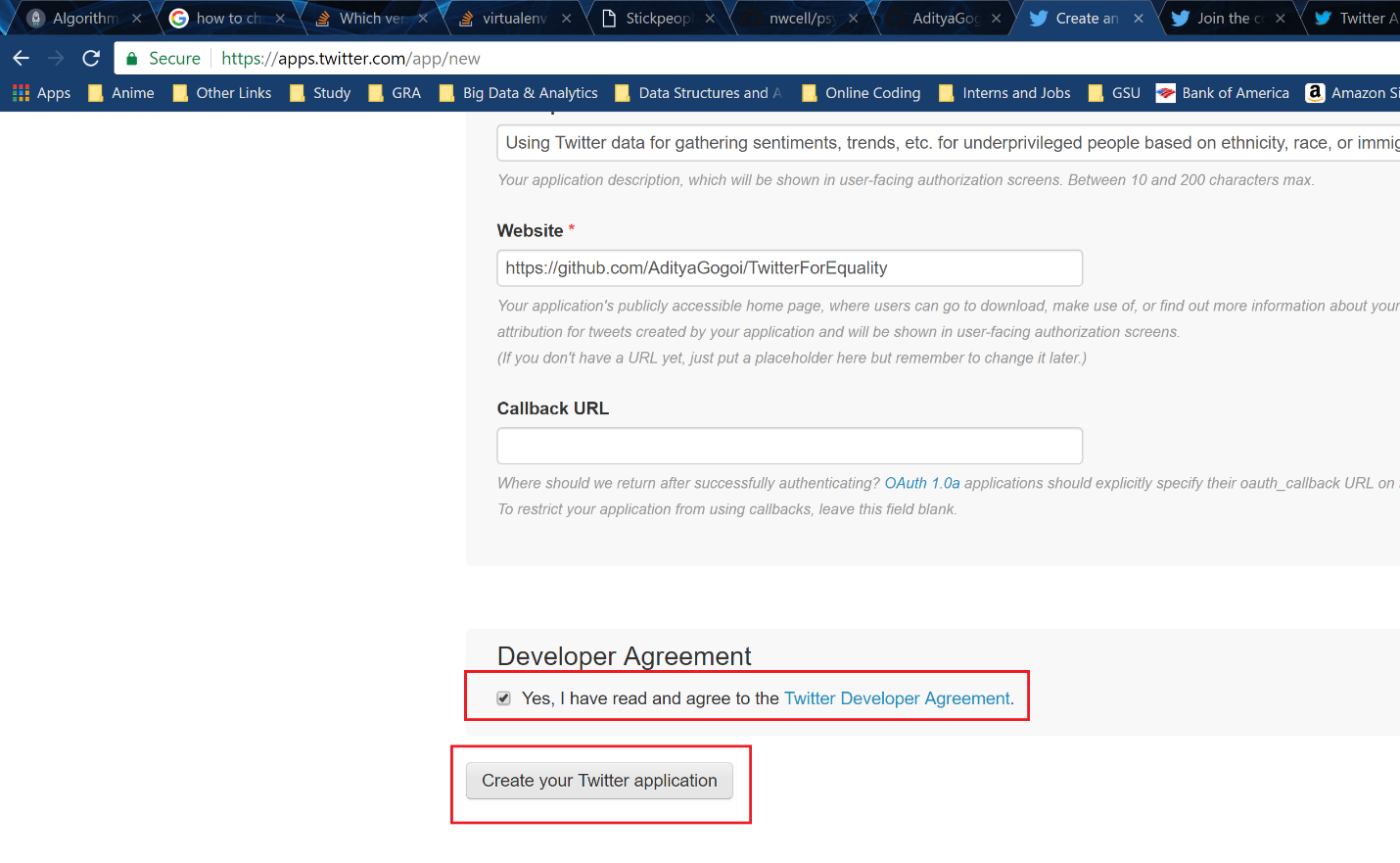
2. You will see a screen saying you do not have any apps yet. Click on the “Create a new app” button.



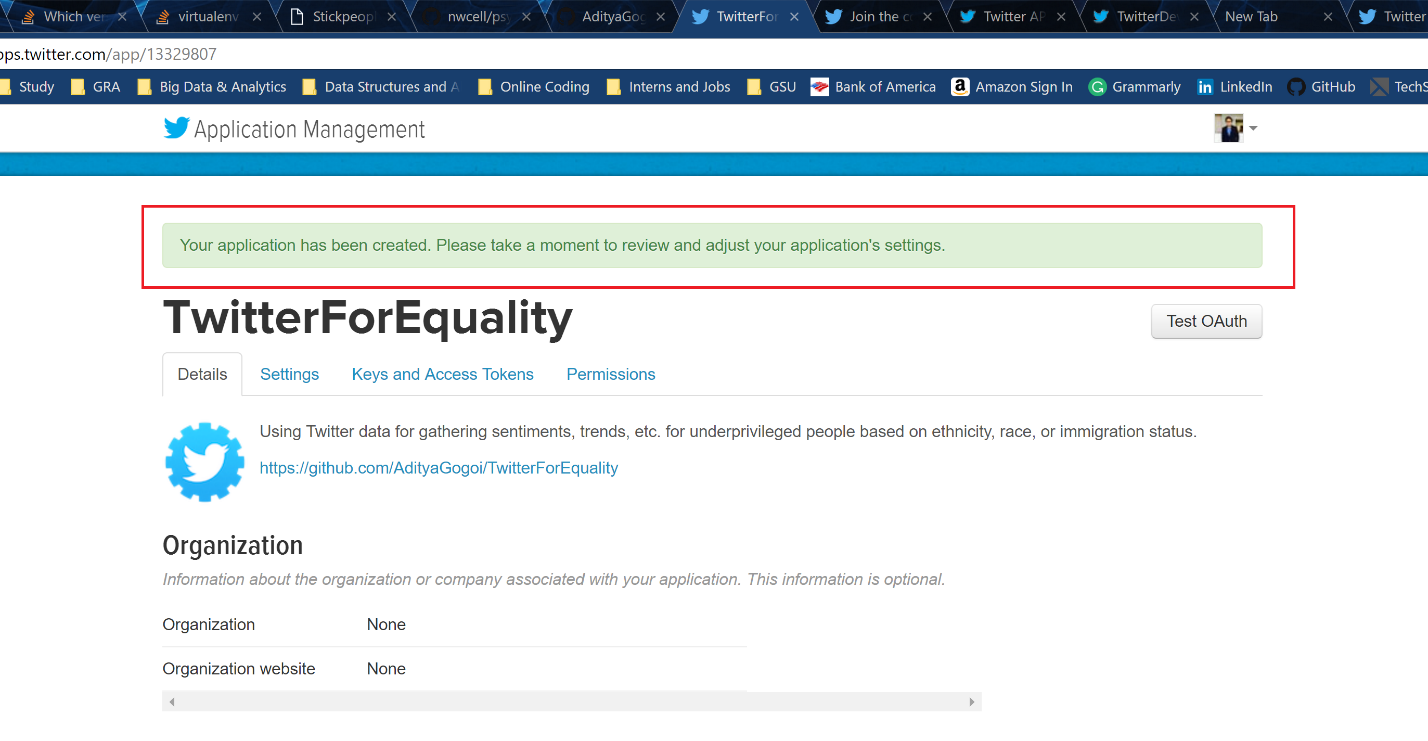
3. Provide Details for the App. I have kept the Website for the app as my Github repository for the same project.



4. In the same page below, check the Terms and Conditions statement and click Create button. Note:- You should have all the necessary details like Mobile Number,etc. saved to your Twitter profile before you create the app.



5. You will get a success message. This page contains important information. Do not close it yet.



**Getting Python ready for the code:-**

Before running the program, we have to install the libraries that python will use for running the code. On running these commands, you will either get an “Installation Success” or “Requirement Already Satisfied” message. Both are fine.

1. Open command prompt and run the following commands as they are without any changes:-

pip install tweepy

pip install ipython

pip install scipy

pip install numpy

pip install pandas

pip install numpy

pip install dataset

pip install textblob

(Installing psycogp2 is tricky because it does not have a version-free, direct pip installation link yet. We can still use this Github link code to install it based on the Python version you have installed. I have included both the 3.x and 2.x installation commands. You can run both and see what works.)

(For Python 3.x )

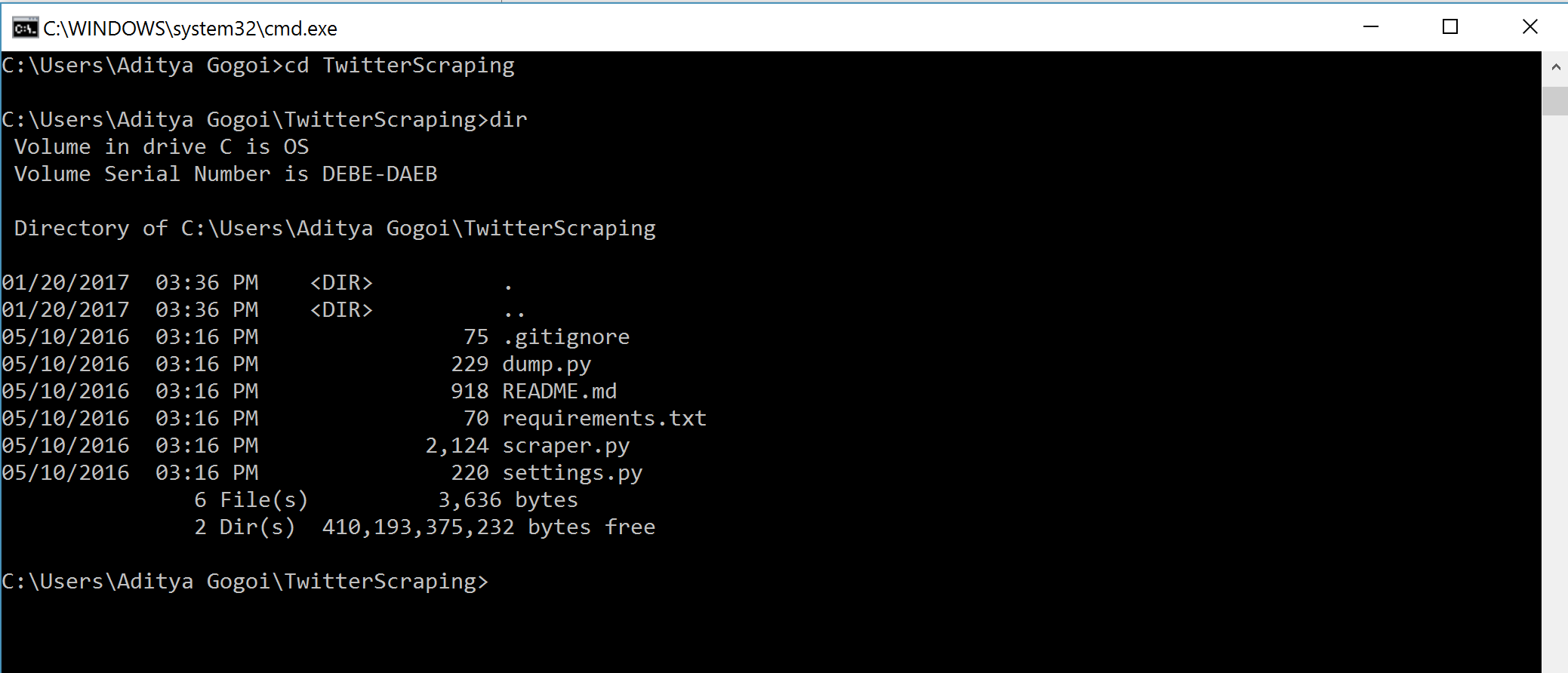
pip install git+https://github.com/nwcell/psycopg2-windows.git@win64-py34#egg=psycopg2

(For Python 2.x)

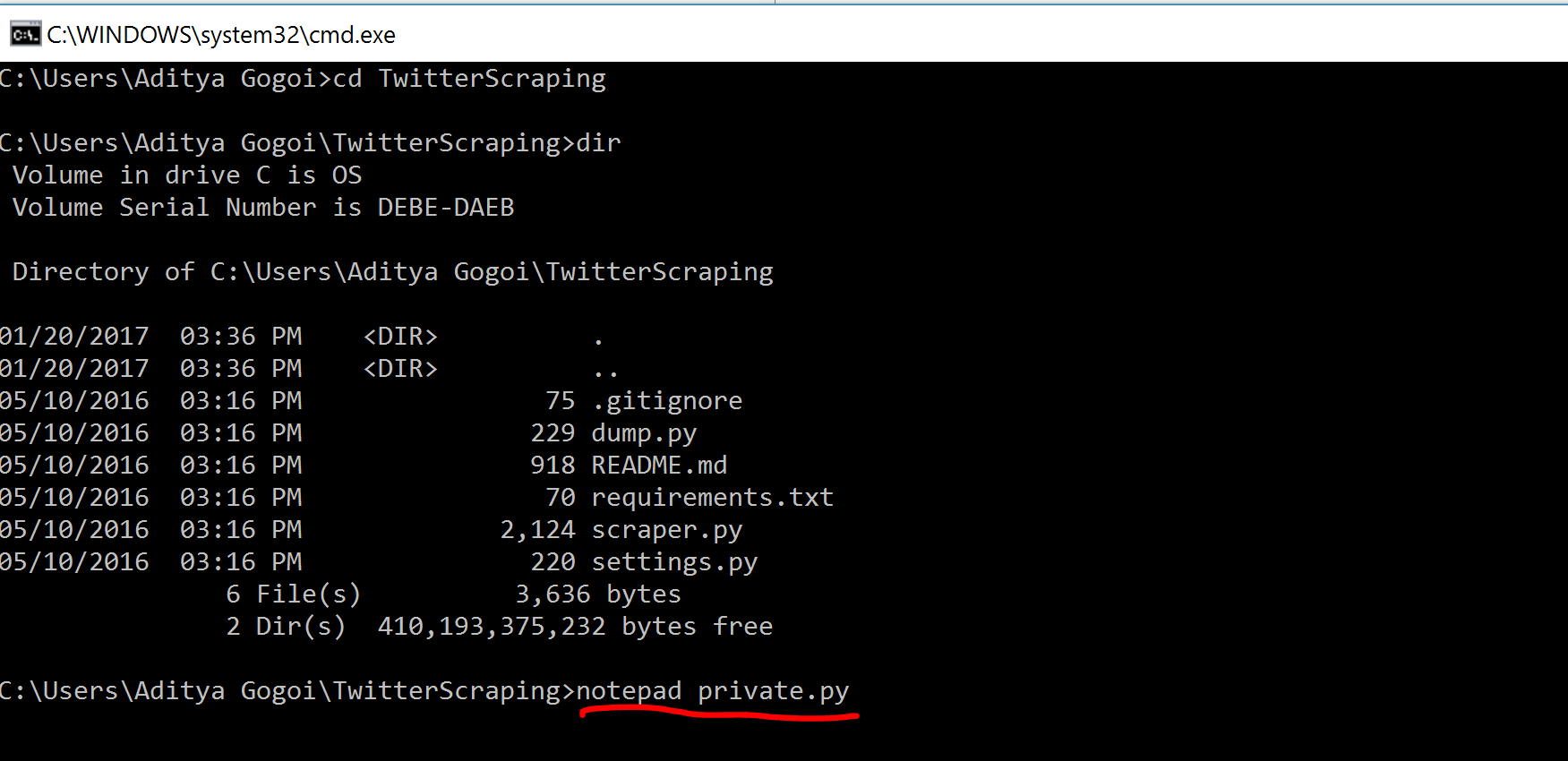
pip install git+https://github.com/nwcell/psycopg2-windows.git@win64-py27#egg=psycopg2

**Configuring the code:-**

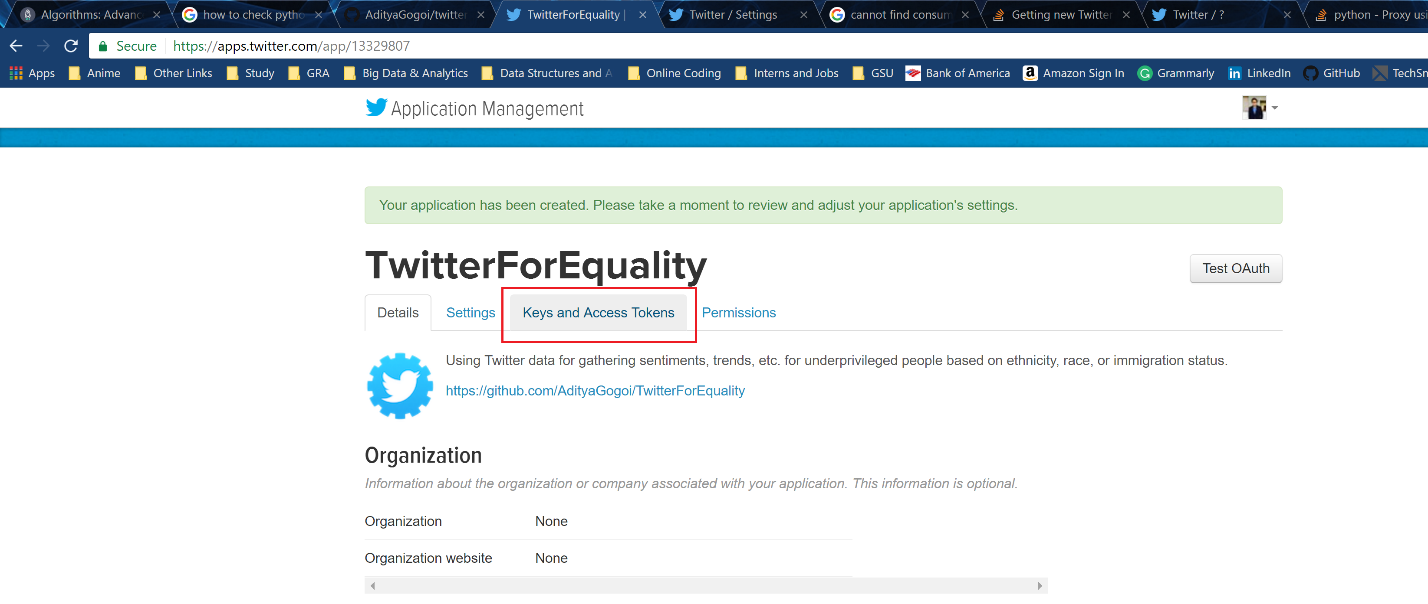
1. Open command prompt and open the folder containing the codes. For ease of access, copy and paste your program folder to the default directory of Command Prompt.



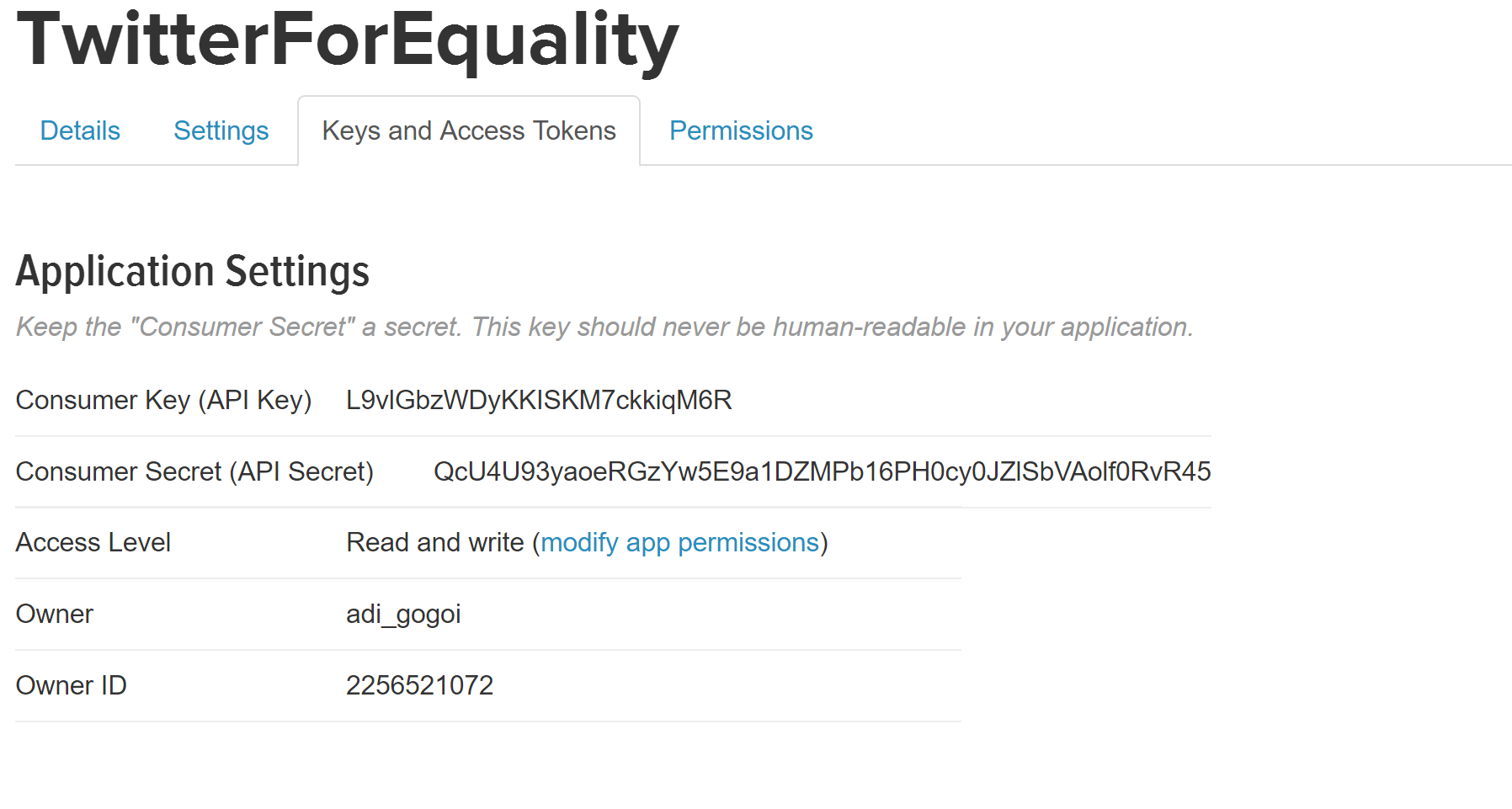
2. Create a python file private.py with the help of notepad. You will get a message “private.py does not exist. Do you want to create it?”. Click Yes.



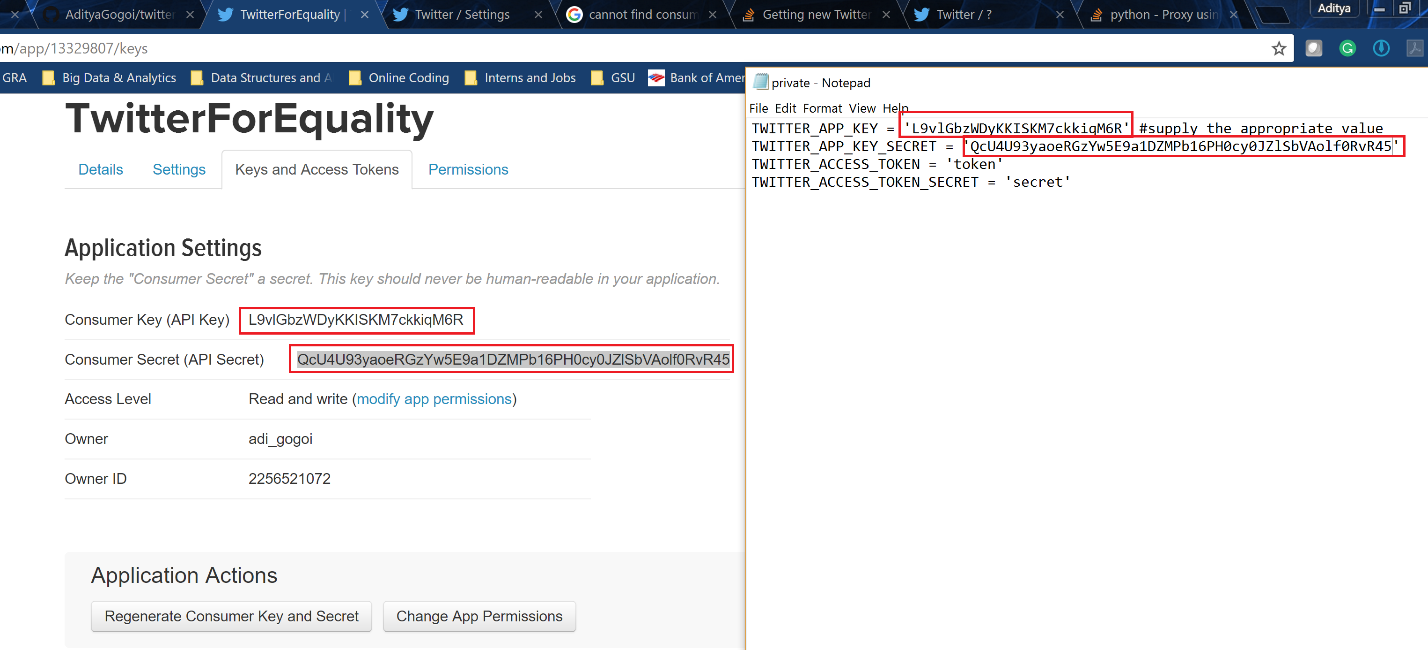
3. In the page where we had successfully created the application click on the “Keys and Access Tokens” tab.



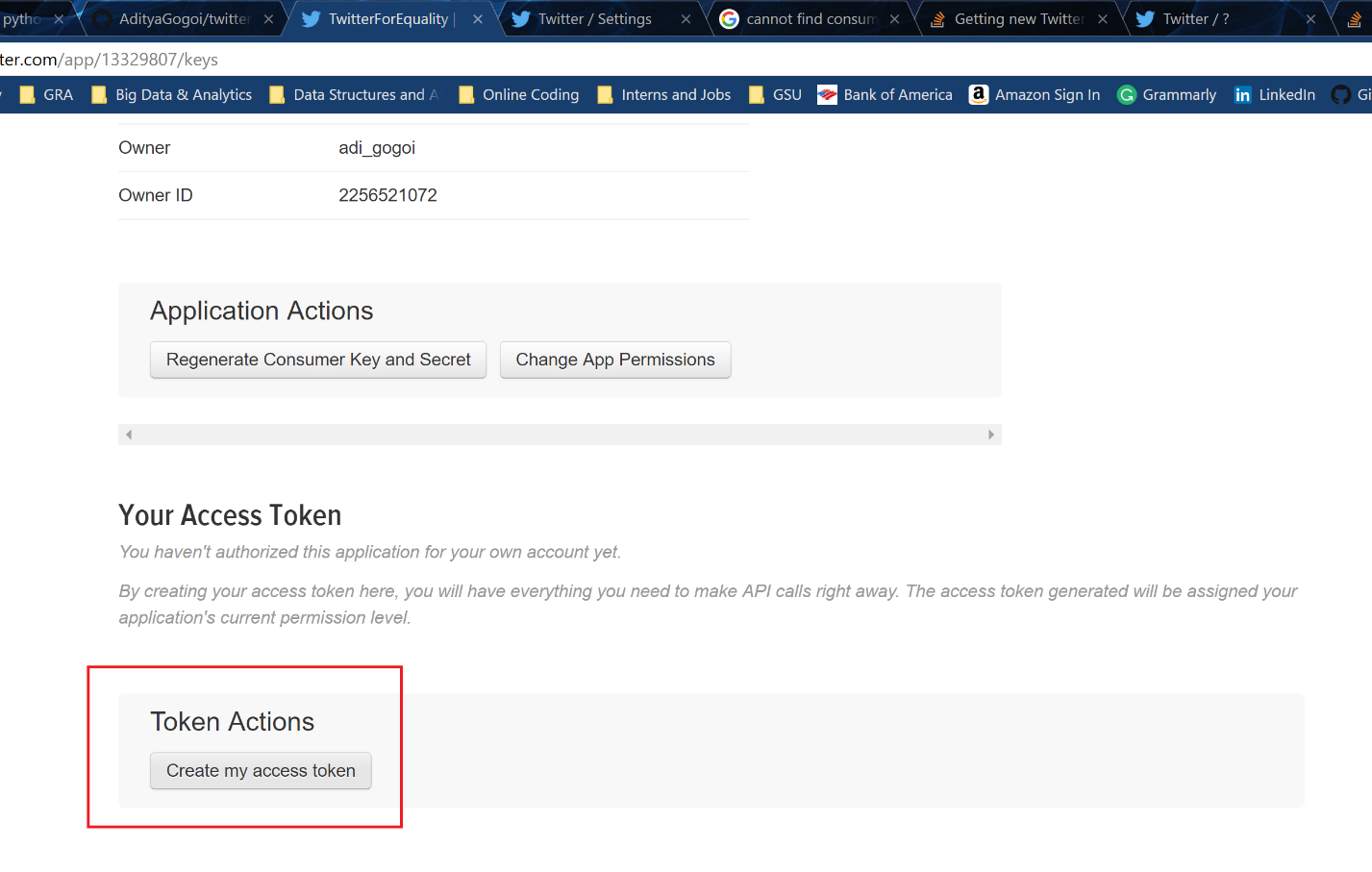
4. In the tab, we can see many tokens. We will use these on the private.py for settings.



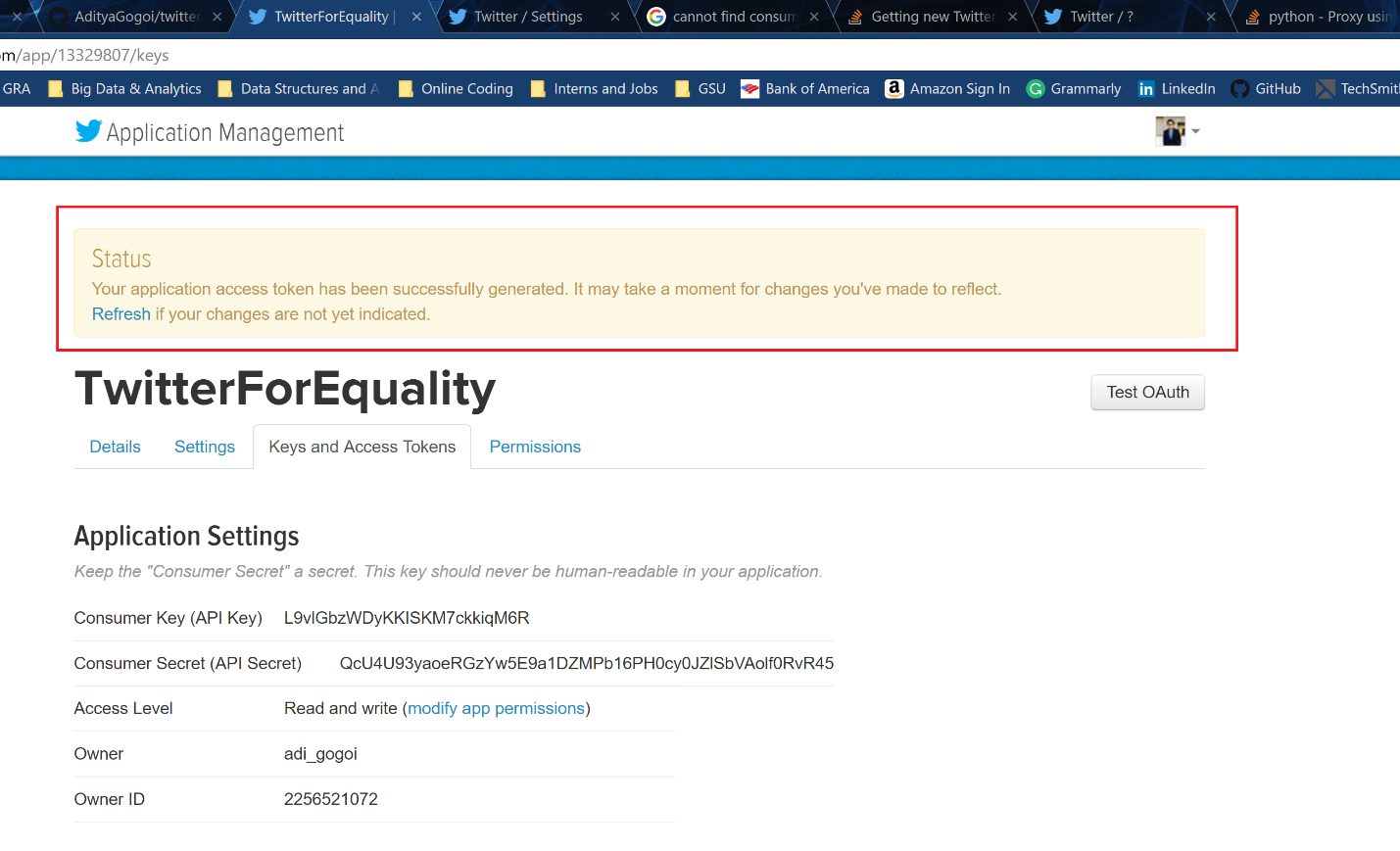
5. Add the Consumer Key and Secret on the fields TWITTER\_APP\_KEY and TWITTER\_APP\_KEY\_SECRET respectively.



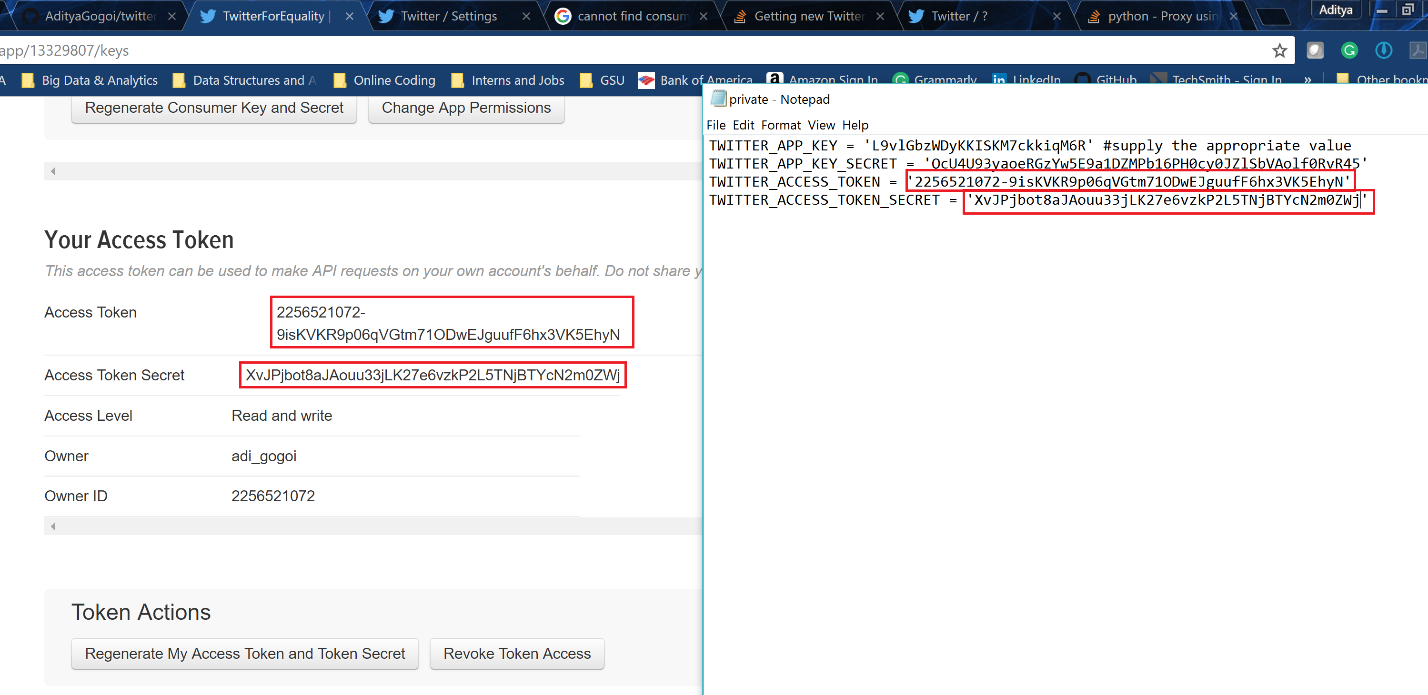
6. Go down on the same webpage an click on the ‘Create my access token’.



7. On the next page you will get a success message.

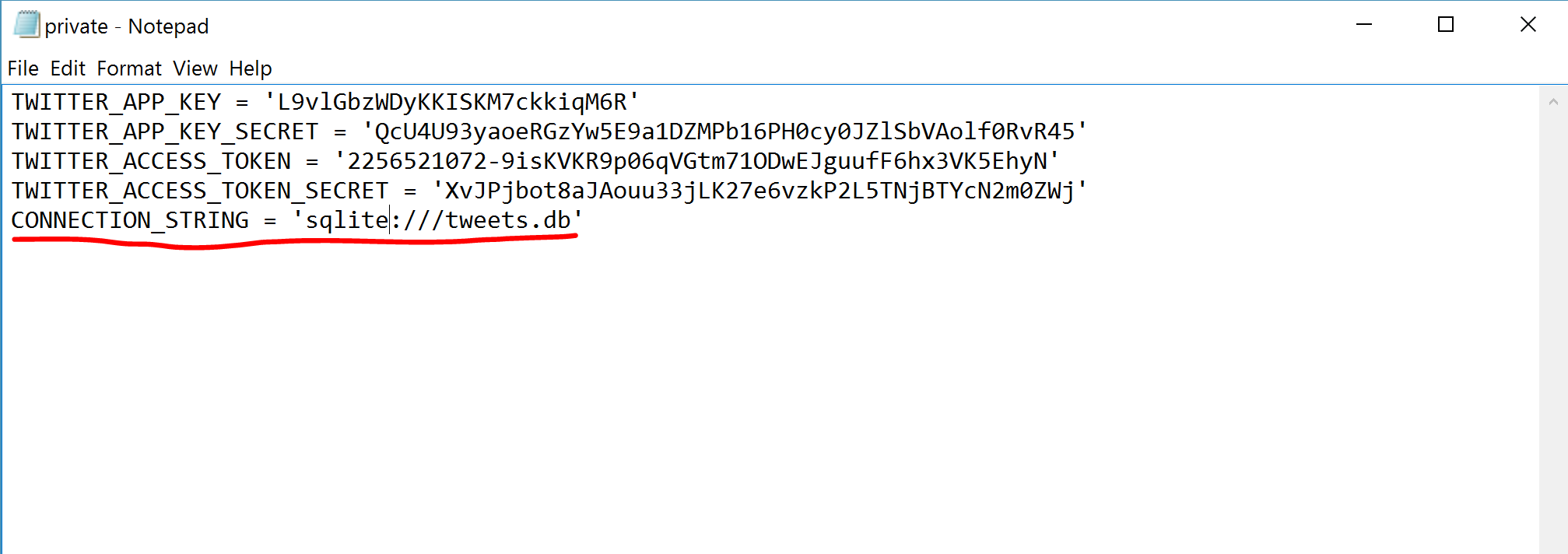


8. Going down on the same page, you will get your AccessToken and Access Token Secret. Copy and paste them to the TWITTER\_ACCESS\_TOKEN and TWITTER\_ACCESS\_TOKEN\_SECRET fields respectively.

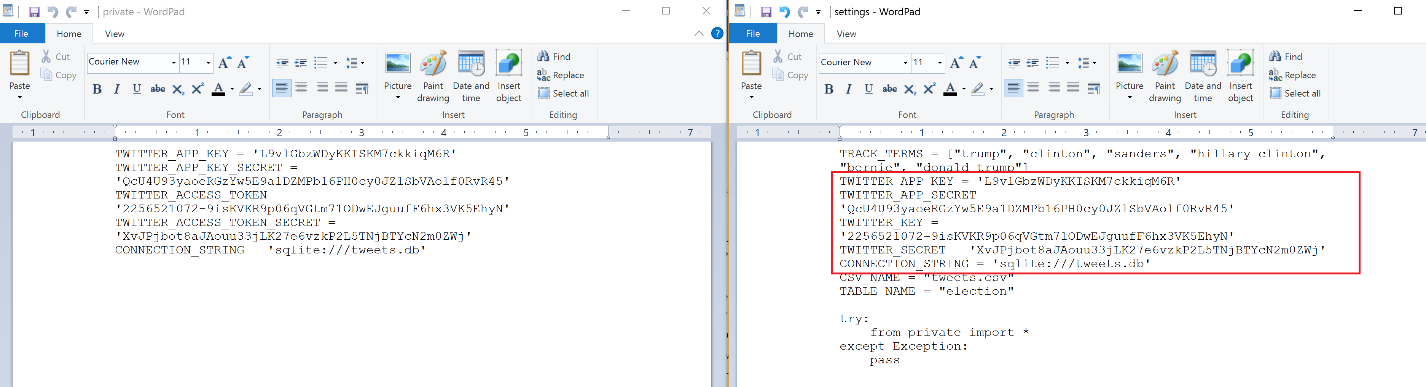


9. Add another configuration to the file.

CONNECTION\_STRING = 'sqlite:///tweets.db'

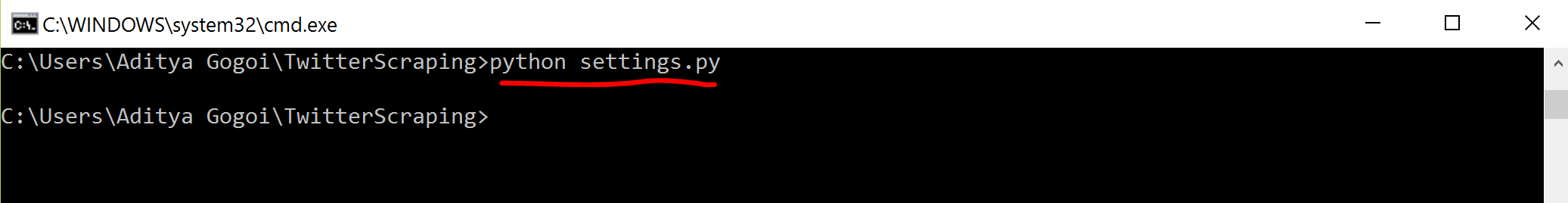


Also add the same data to the settings.py program with slight changes in name. TWITTER\_APP\_KEY\_SECRET will be TWITTER\_APP\_SECRET, TWITTER\_ACCESS\_TOKEN will be TWITTER\_KEY and TWITTER\_ACCESS\_TOKEN\_SECRET will be TWITTER\_SECRET.

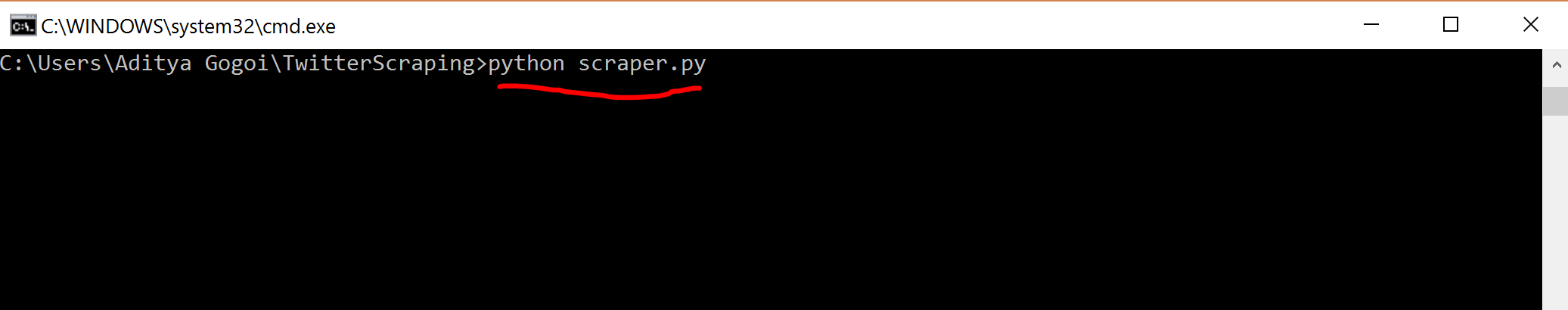


10. Save and close private.py after filling it with above values.

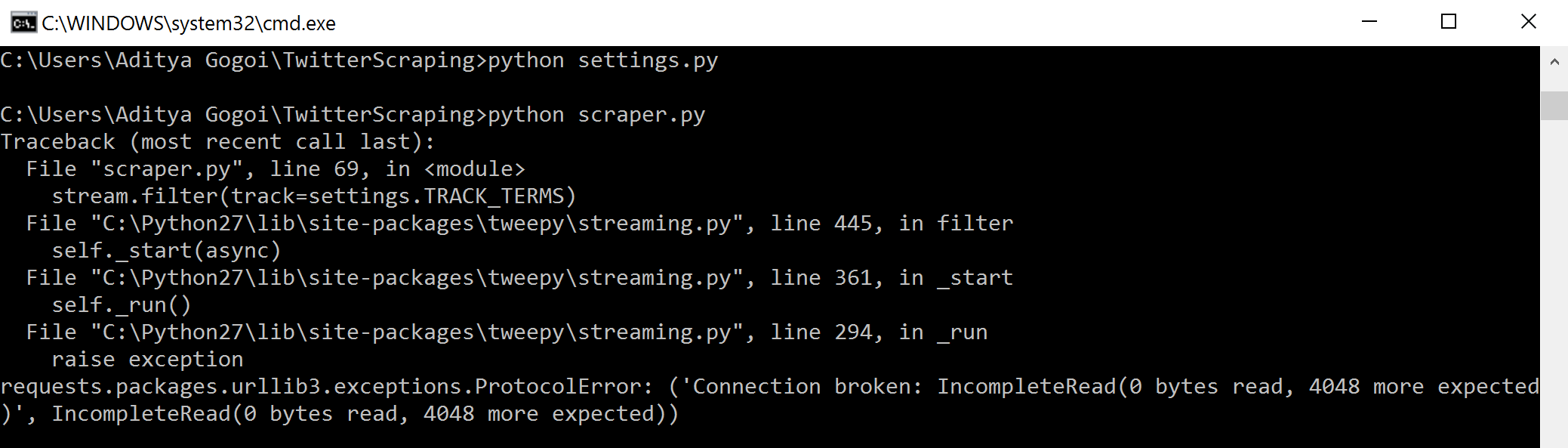
11. Run settings.py for exporting all the configuration data.



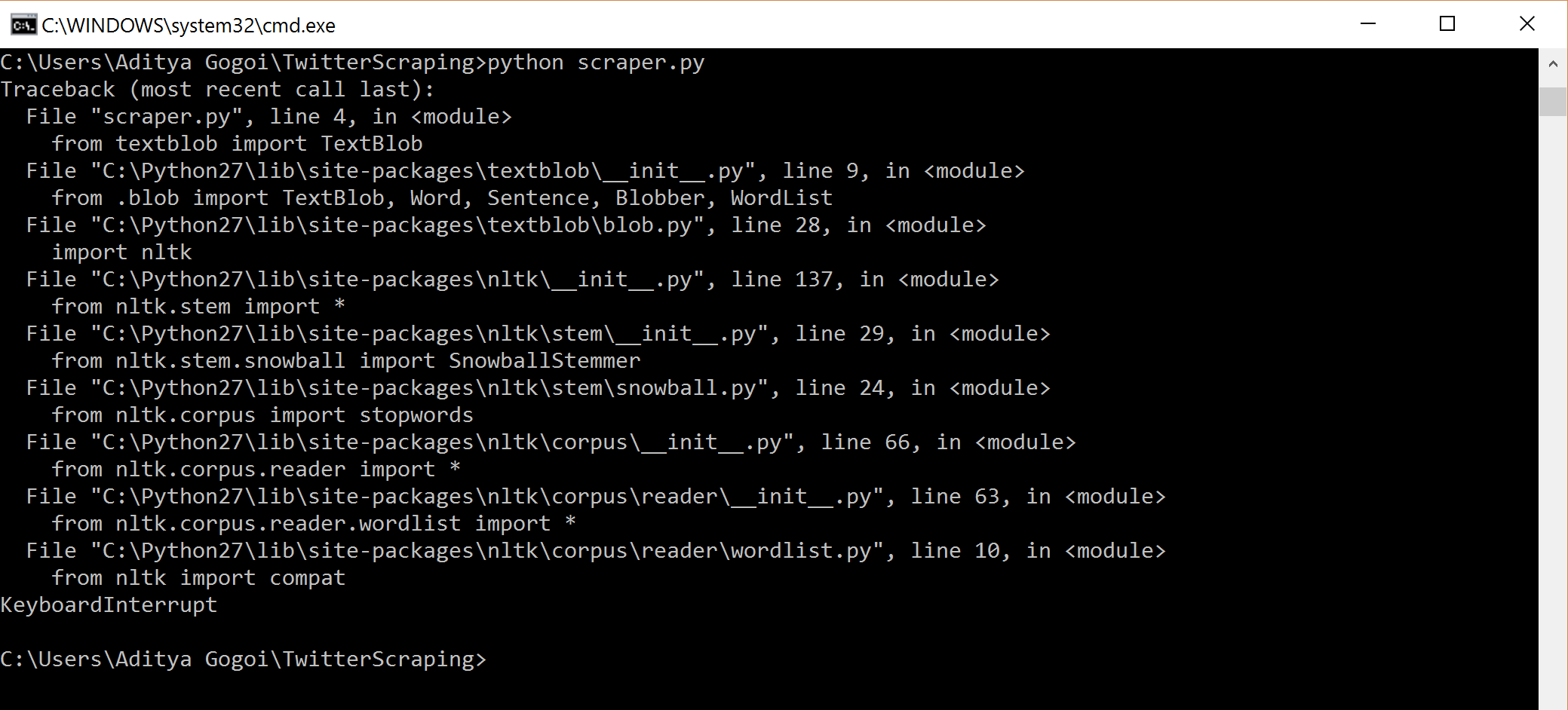
11. Run the scraper.py program to start scraping.



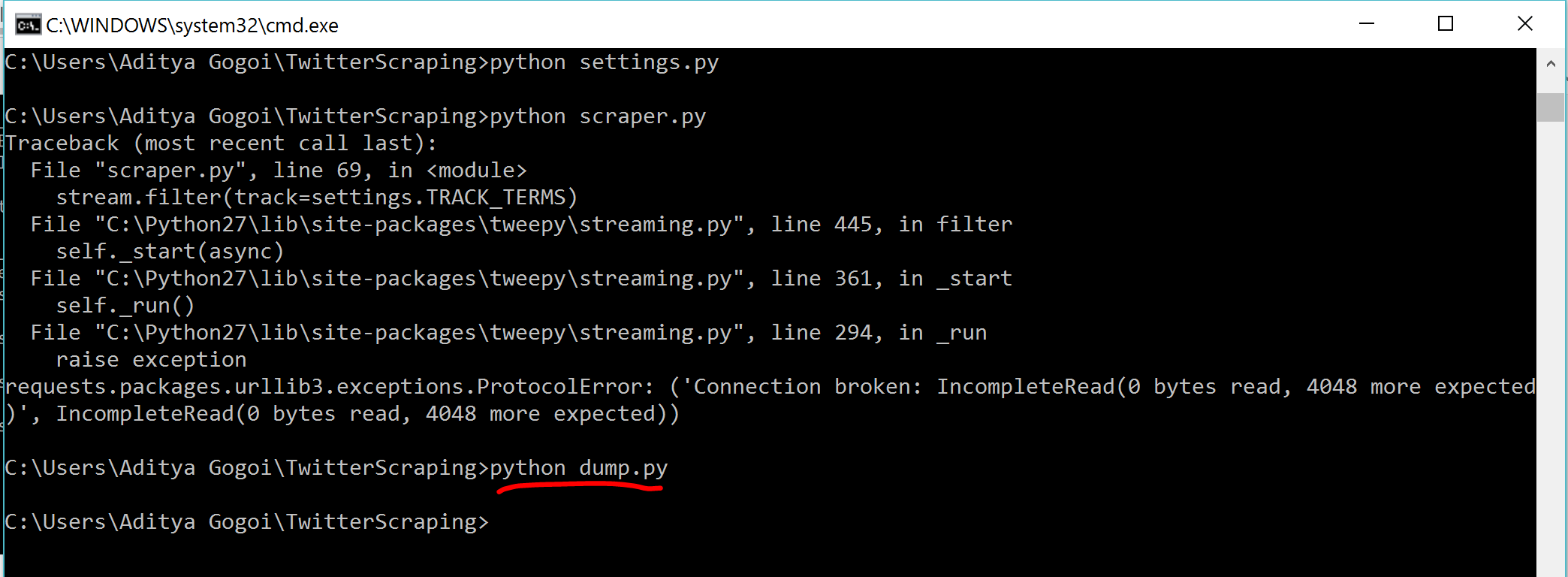
Note:- Sometimes due to connection error, the program might stop unexpectedly like this.



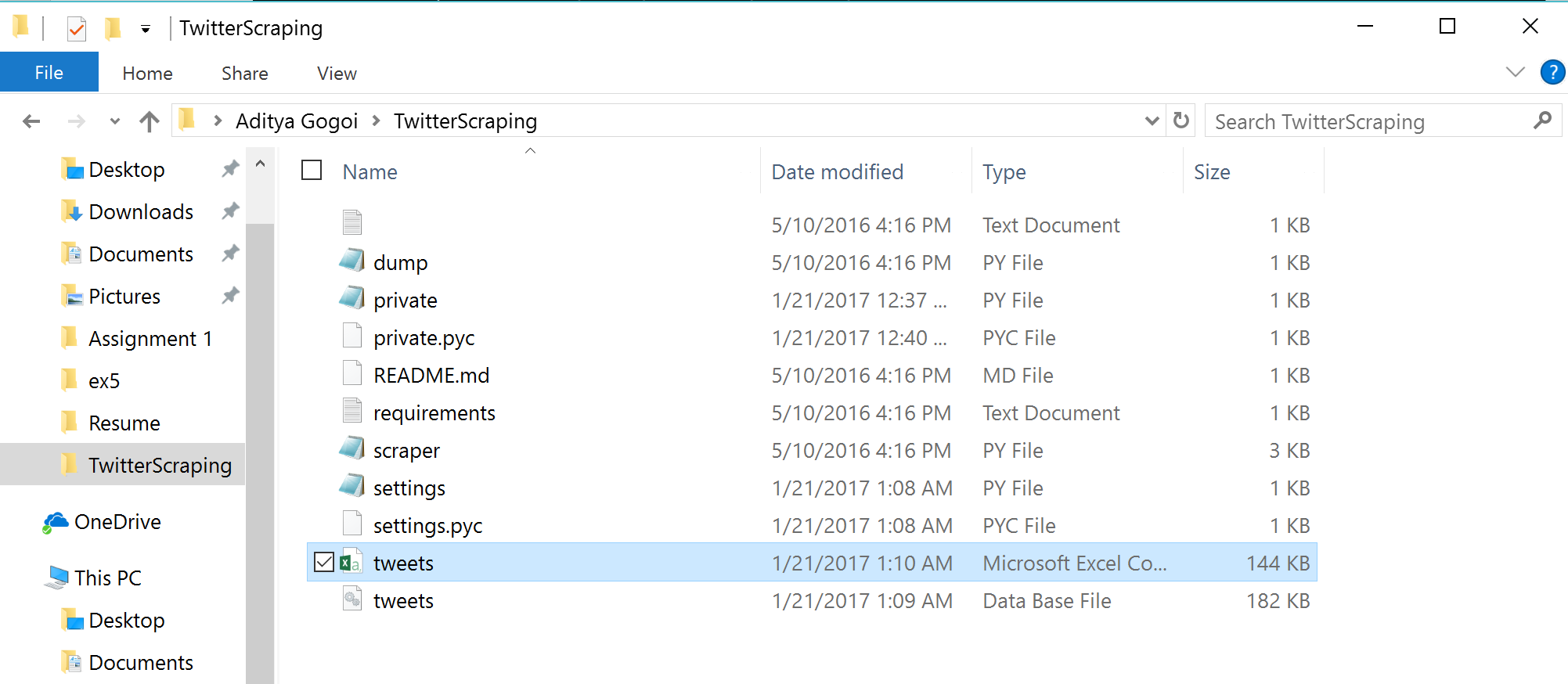
12. Hit ‘Ctrl+c’ to halt the program.



13. Now run the program dump.py to create a generate dump.csv file, which will contain data that was scrapped.



14. In the folder where all the programs are, you will see a new csv file created called tweets.csv. This is the dump that contains the tweets that were scraped.



15. On taking a look at the csv file, we see many interesting data related to the tweets like when it was created, the text in the tweet, user name, user description, when user was created, and tweet subjectivity. The ‘tweet\_subjectivity’ is a field that uses a python library that cleans the tweet and then uses text Blob.sentiment to see the users subjectivity to the tweet.

