Twitter API - run instructions

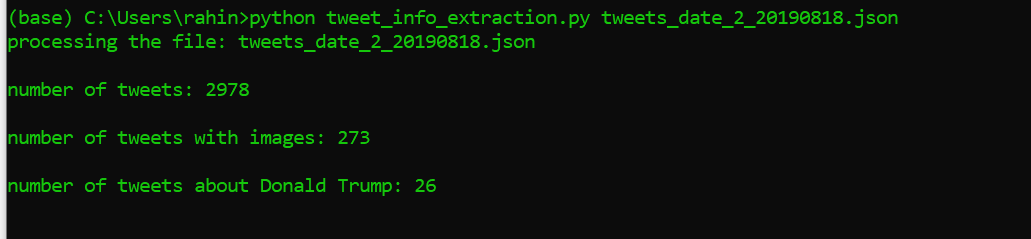
# Step 1: Get twitter dataset in its json format:

* Collect JSON twitter dataset (containing Tweets and Retweets) using the script twitter\_streaming\_api\_complete.py and send output stream to a .json file as seen in the below snippet.
* Note: Per API rules, the script waits for 60minutes whenever the streaming rate limits is reached and resumes from previous checkpoint.

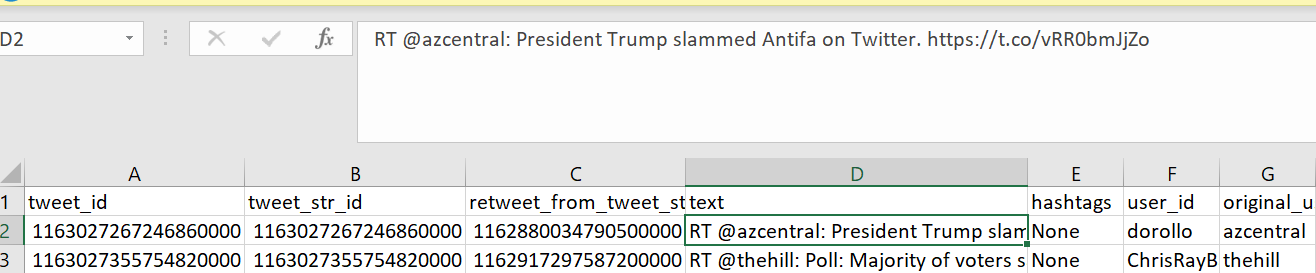


# Step 2: Create .csv file from processing the above .json file:

* Pass the .json file as the argument to the script tweet\_info\_extraction.py. 🡪 this script filters the stream data to retain only those tweets which has (1) images and (2) words ‘donald trump’



* Data in .csv file:

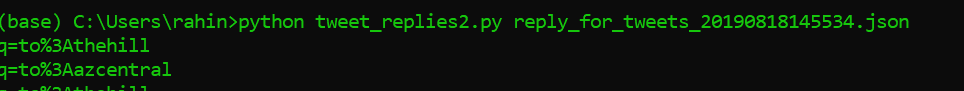


* (1) file with required info to gather replies – input to next step and (2) tweets .csv file:



# Step 3: collecting replies:

Finally, run the below script with the .json file created from previous script as argument.



The below script creates two files of replies: (1) recent replies to all posts of the newspaper handle and (2) replies to posts on trump. Since the api search can return anywhere between 15 to 100 replies, the no. (2) file may not be populated all the time. Therefore repeated running is required.



