**Twitter Scraping Using Streamlit and Snscrape**

|  |  |
| --- | --- |
| Project Title | Twitter Scraping |
| Skills take away From This Project | Python scripting, Data Collection, MongoDB, Streamlit |
| Domain | Social Media |

**Problem Statement:**

Today, data is scattered everywhere in the world. Especially in social media, there may be a big quantity of data on Facebook, Instagram, Youtube, Twitter, etc. This consists of pictures and films on YouTube and Instagram as compared to Facebook and Twitter. To get the real facts on Twitter, you want to scrape the data from Twitter. You Need to Scrape the data like (date, id, url, tweet content, user, reply count, retweet count, language, source, like count etc)from twitter.

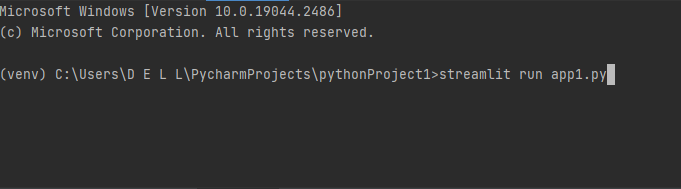
**Approach**:

* By using the “**snscrape**” Library, Scrape the twitter data from Twitter [Reference](https://medium.com/dataseries/how-to-scrape-millions-of-tweets-using-snscrape-195ee3594721)
* **Create a dataframe with** date, id, url, tweet content, user,reply count, retweet count, language, source, like count.
* **Store each collection of data** into a document into Mongodb along with the hashtag or key word we use to  Scrape from twitter. eg:({“LIC corporation+current Timestamp”: [{1000  Scraped data from past 100 days }]})
* **Create a GUI using** [**streamlit**](https://docs.streamlit.io/library/api-reference) that should contain the feature to enter the keyword or Hashtag to be searched, select the date range and limit the tweet count need to be scraped. After scraping, the data needs to be displayed in the page and need a button to upload the data into Database and download the data into csv and json format.

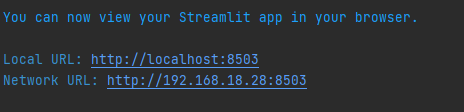
Step 1:

Install the packages pandas, streamlit, pymongo, snscrape and pycharm version 2020.3 before executing the project;

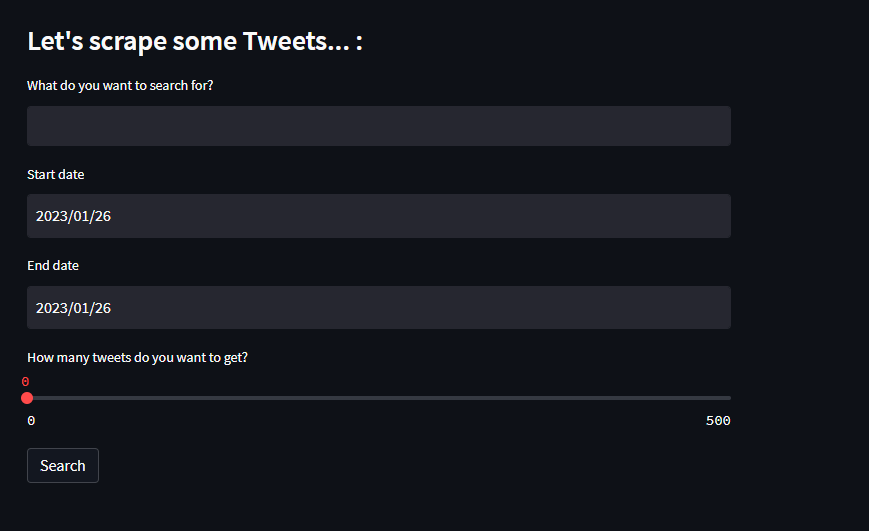
Now, Run the program app1.py in pycharm using streamlit command.



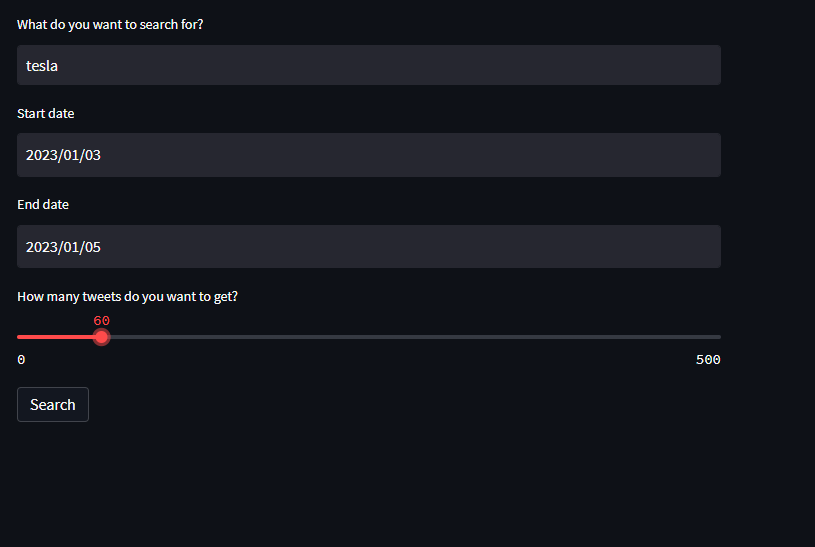
Step2: You will get link that opens in browser



Step 3: The Screen appears like this

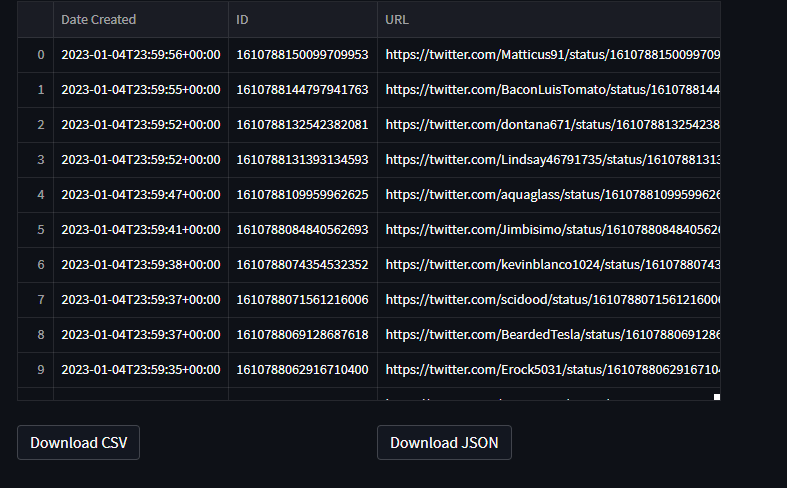


Step 4: Give the input details



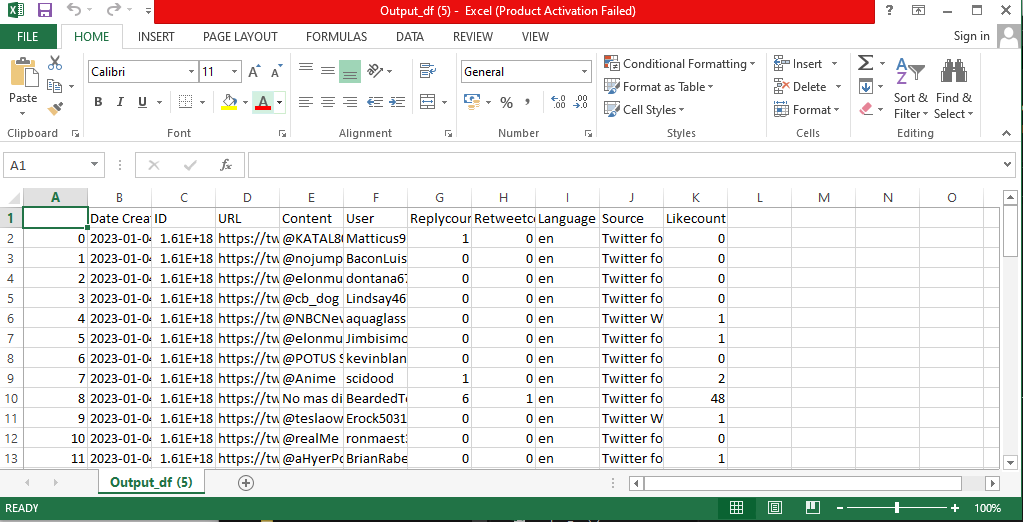
Step 5:

The dataframe appears as below



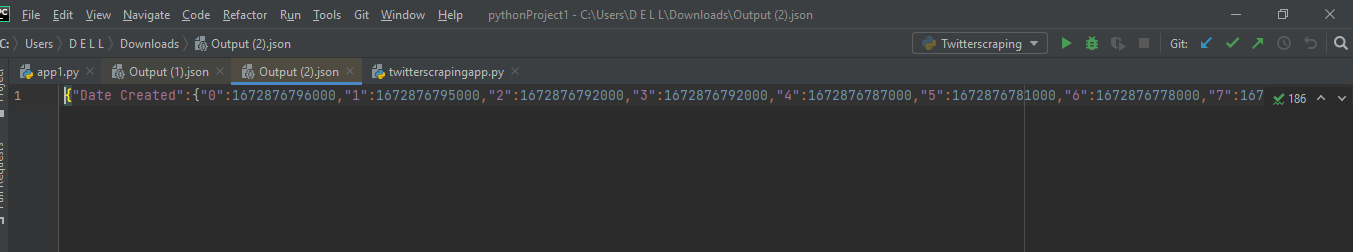
Step 6:

Now you can download the file either in csv or in json format, the csv file is as below



Step 7:

The json file is as below



Step 8:

The same dataframe gets updated in Cloud .(MongoDB Atlas)

