**Steps for Tweets Extraction**

1. To extract tweets from Twitter, you need to Download Tweepy
2. To get tweets from Twitter, you need 4 things: consumer\_key,consumer\_secret,access\_token,access\_token\_secret

Link to get this information

https://developer.twitter.com/en/portal/projects-and-apps

https://www.slickremix.com/docs/how-to-get-api-keys-and-tokens-for-twitter/

Get your credential and fill it to use this code

consumer\_key= '.................'

consumer\_secret= '.....................'

access\_token= '......................'

access\_token\_secret= '...........................'

1. Code to connect the Twitter using above information

auth = tw.OAuthHandler(consumer\_key, consumer\_secret)

auth.set\_access\_token(access\_token, access\_token\_secret)

api = tw.API(auth, wait\_on\_rate\_limit=True)

1. # Collect tweets

tweets = tw.Cursor(api.search,

q=search\_words, ##query or tweets will be retrieved based words given to search\_words

lang="en", ##it will retrieve tweets written in English language

since=date\_since, ## it will retrieve tweets from this date

include\_rts=False, ## exclude re-tweets

exclude\_replies=True, ## exclude replies

contributor\_details=False, ## exclude contributor’s details

tweet\_mode = 'extended' ## tweet\_mode = 'extended' for extended tweets(280 characters) otherwise it will consider only 140 characters.

).items(3000) ## number of records you want to retreive

Example for search\_words

Example1

#search\_words = "@WorldOceansDay" ###

Example 2

search\_words = '(@WorldOceansDay OR World Ocean Day OR #WorldOceanDay) AND (cloud OR plastic OR climate) '

1. Get information from the tweets: You can get Tweet id, location of user, user name, tweets and many more. Visit <https://towardsdatascience.com/how-to-scrape-more-information-from-tweets-on-twitter-44fd540b8a1f> for more information

tweet.id,

tweet.user.screen\_name,

tweet.user.location,

tweet\_list

1. Convert the list to dataframe

tweet\_text = pd.DataFrame(data=users\_locs, columns=['Tweet Id','user', "location", "Tweets"])

1. Save the file to csv format

tweet\_text.to\_csv('tweets\_Ocean.csv')

\*\*\* dataset will be ready with name 'tweets\_Ocean.csv’

\*\*\*\* dataset is related to the tweets posted by the user on the World Ocean Day which was on June 8, 2021