

```
In [10]: #You need to create a twitter developer account
#pip install tweepy
#conda install -c conda-forge textblob
#python -m textblob.download_corpora
```

```
In [11]: #Example from Just added more notes
#https://www.earthdatascience.org/courses/use-data-open-source-python/intro
```

### Authorizing an application to access Twitter account data

- To access the Twitter API, you will need 4 things from the your Twitter App page.
- These keys are located in your Twitter app settings in the Keys and Access Tokens tab.

```
In [12]: #import libraries that you need
import tweepy
from tweepy import OAuthHandler, Stream

#import other required libraries
import os
import pandas as pd

from credentials import *
```

- Go to <https://developer.twitter.com/en/apps> (<https://developer.twitter.com/en/apps>) to create an app and get values
- for these credentials, which you'll need to provide in place of these
- empty string values that are defined as placeholders.
- See <https://developer.twitter.com/en/docs/basics/authentication/overview/oauth> (<https://developer.twitter.com/en/docs/basics/authentication/overview/oauth>)
- for more information on Twitter's OAuth implementation.

```
In [14]: #create a file called credentials.py make sure it is in the same folder as
# the credentials file will look like this
#ACCESS_TOKEN = 'xxx'
#ACCESS_SECRET = 'xx'
#CONSUMER_KEY = 'xx'
#CONSUMER_SECRET = 'xx'
```

In [13]:

```
auth = tweepy.OAuthHandler(CONSUMER_KEY, CONSUMER_SECRET)
auth.set_access_token(ACCESS_TOKEN, ACCESS_SECRET)

twitter_api = tweepy.API(auth)

# Nothing to see by displaying twitter_api except that it's now a
# defined variable

print(twitter_api)
```

<tweepy.api.API object at 0x7fa898bfbe50>

### Search Twitter for Tweets

- Now you are ready to search Twitter for recent tweets!
- Start by finding recent tweets that use the #wildfires hashtag.
- You will use the .Cursor method to get an object containing tweets containing the hashtag #wildfires.
- To create this query, you will define the:
  - Search term - in this case #wildfires
  - the start date of your search
  - Remember that the Twitter API only allows you to access the past few weeks of tweets, so you cannot dig into the history too far.

In [13]:

```
# Define the search term and the date_since date as variables
search_words = "#wildfires"
date_since = "2018-11-16"
```

In [15]:

```
#Below you use .Cursor() to search twitter for tweets containing the search
#You can restrict the number of tweets returned by specifying a number in t
#.items(5) will return 5 of the most recent tweets.
```

```
#Collect tweets
tweets = tweepy.Cursor(twitter_api.search_tweets,
                        q=search_words,
                        lang="en",
                        since_id=date_since).items(5)

tweets
```

Out[15]: <tweepy.cursor.ItemIterator at 0x7fb8c8147c40>


- .Cursor() returns an object that you can iterate or loop over to access the data collected.
- Each item in the iterator has various attributes that you can access to get information about each tweet including:
  - the text of the tweet
  - who sent the tweet


- the date the tweet was sent and more
- The code below loops through the object and prints the text associated with each tweet.


```
In [17]: # Collect tweets
tweets = tweepy.Cursor(twitter_api.search_tweets,
                        q=search_words,
                        lang="en",
                        since_id=date_since).items(5)

# Iterate and print tweets
for tweet in tweets:
    print(tweet.text)
```

wildfire in Ruidoso, New Mexico. More than 150 houses burned down as a result of this forest fire. #USA #wildfire... <https://t.co/cUkUB7Cjab> (<https://t.co/cUkUB7Cjab>)

RT @TessaMentus: #NM #Wildfires Latest   
@KOB4

#McBrideFire in @VillageRuidoso  
5300+ acres, 0% containment, 150+ structures gone  
Cause: u...  
RT @TessaMentus: #NM #Wildfires Latest   
@KOB4

#McBrideFire in @VillageRuidoso  
5300+ acres, 0% containment, 150+ structures gone  
Cause: u...  
#NM #Wildfires Latest   
@KOB4

#McBrideFire in @VillageRuidoso  
5300+ acres, 0% containment, 150+ structures gone  
Cause: under investigation  
RT @SVB\_Musician: Just went to the gas station to get some sodas and bottled waters. This is really really bad! #McBrideFire burning in Rui...

- The above approach uses a standard for loop.
- However, this is an excellent place to use a Python list comprehension.
- A list comprehension provides an efficient way to collect object elements contained within an iterator as a list.

```
In [19]: # Collect tweets
tweets = tweepy.Cursor(twitter_api.search_tweets,
                        q=search_words,
                        lang="en",
                        since_id=date_since).items(5)

# Collect a list of tweets
[tweet.text for tweet in tweets]
```

```
Out[19]: ['wildfire in Ruidoso, New Mexico. More than 150 houses burned down as a
result of this forest fire. #USA #wildfire... https://t.co/cUkUB7Cjab', (ht
tps://t.co/cUkUB7Cjab'),
'RT @TessaMentus: #NM #Wildfires Latest 🇺🇸 \n@KOB4 \n\n#McBrideFire in @Vi
llageRuidoso \n5300+ acres, 0% containment, 150+ structures gone\nCause:
u...',
'RT @TessaMentus: #NM #Wildfires Latest 🇺🇸 \n@KOB4 \n\n#McBrideFire in @Vi
llageRuidoso \n5300+ acres, 0% containment, 150+ structures gone\nCause:
u...',
'#NM #Wildfires Latest 🇺🇸 \n@KOB4 \n\n#McBrideFire in @VillageRuidoso \n53
00+ acres, 0% containment, 150+ structures gone\nCause: under investigati
on',
'RT @SVB_Musician: Just went to the gas station to get some sodas and bo
ttled waters. This is really really bad! #McBrideFire burning in Rui...']
```

### To Keep or Remove Retweets

- A retweet is when someone shares someone else's tweet.
- It is similar to sharing in Facebook.
- Sometimes you may want to remove retweets as they contain duplicate content that might skew your analysis if you are only looking at word frequency.
- Other times, you may want to keep retweets.
- Below you ignore all retweets by adding `-filter:retweets` to your query.
- The Twitter API documentation has information on other ways to customize your queries.

```
In [20]: new_search = search_words + " -filter:retweets"
new_search
```

```
Out[20]: '#wildfires -filter:retweets'
```

```
In [21]: tweets = tweepy.Cursor(twitter_api.search_tweets,
                                q=new_search,
                                lang="en",
                                since_id=date_since).items(5)

[tweet.text for tweet in tweets]
```

```
Out[21]: ['wildfire in Ruidoso, New Mexico. More than 150 houses burned down as a
result of this forest fire. #USA #wildfire... https://t.co/cUkUB7Cjab', (ht
tps://t.co/cUkUB7Cjab',)
'#NM #Wildfires Latest 🇺🇸 \n@KOB4 \n\n#McBrideFire in @VillageRuidoso \n3
00+ acres, 0% containment, 150+ structures gone\nCause: under investigati
on',
"@BeantownCanuck @arthister AND--all deaths related to BC's EXTREME weat
her events last 3 years- #wildfires #floods... https://t.co/ty7zDv5uhJ", (h
ttps://t.co/ty7zDv5uhJ',)
'Using donations, our #organization has responded to #wildfires disaster
s to provide essentials for survivors. 📦 \n\nPl... https://t.co/I0UIKwIdn
3', (https://t.co/I0UIKwIdn3',)
'#sealevel #drought #Wildfires #climaterefugee #climaterefugees #Web3 #N
FT #domainname #ClimateCrisis #GlobalWarming... https://t.co/uFUK2iA9Nn']
(https://t.co/uFUK2iA9Nn']])
```

### Who is Tweeting About Wildfires?

- You can access a wealth of information associated with each tweet. Below is an example of accessing the users who are sending the tweets related to #wildfires and their locations.
- Note that user locations are manually entered into Twitter by the user.
- Thus, you will see a lot of variation in the format of this value.
  - tweet.user.screen\_name provides the user's twitter handle associated with each tweet.
  - tweet.user.location provides the user's provided location.
- You can experiment with other items available within each tweet by typing tweet. and using the tab button to see all of the available attributes stored.

```
In [22]: tweets = tweepy.Cursor(twitter_api.search_tweets,
                                q=new_search,
                                lang="en",
                                since_id=date_since).items(5)

users_locs = [[tweet.user.screen_name, tweet.user.location] for tweet in tw
users_locs
```

```
Out[22]: [['Brave_spirit81', 'Ukraine 🇺🇦'],
['TessaMentus', 'Albuquerque'],
['savoirfaire_2', 'Abbotsford BC (Stó:lō Nation)'],
['LiFrAOrg', 'Maxwell, CA'],
['MarijuanaName', '#Mars']]
```

### Create a Pandas Dataframe From A List of Tweet Data

Once you have a list of items that you wish to work with, you can create a pandas dataframe that contains that data.

```
In [23]: tweet_text = pd.DataFrame(data=users_locs,
                                   columns=['user', "location"])
tweet_text
```

Out[23]:

	user	location
0	Brave_spirit81	Ukraine 🇺🇦
1	TessaMentus	Albuquerque
2	savoirfaire_2	Abbotsford BC (Stó:lō Nation)
3	LiFrAOrg	Maxwell, CA
4	MarijuanaName	#Mars

### Customizing Twitter Queries

- For instance, if you search for climate+change, Twitter will return all tweets that contain both of those words (in a row) in each tweet.

```
In [25]: #Note that the code below creates a list that can be queried
#using Python indexing to return the first five tweets.
```

```
new_search = "climate+change -filter:retweets"

tweets = tweepy.Cursor(twitter_api.search_tweets,
                        q=new_search,
                        lang="en",
                        since_id='2018-04-23').items(1000)

all_tweets = [tweet.text for tweet in tweets]
all_tweets[:5]
```

```
Out[25]: ['@alexisleclezio Spot on! But nooo, there'll be another 500 billion loa
n taken then stolen to stop 'climate change'... https://t.co/R2tLLwJsYo',
(https://t.co/R2tLLwJsYo),
'Gorgeous, gorgeous girls hate climate change and wish that this shit wo
uldn't end out planet in like 5 yrs😭',
'@RadioFreeTom Canceling student debt and climate Change are widely popu
lar. What's the problem?',
'@EgoEire @barrywalsh9 @oconnellhugh https://t.co/xD07uUAxVj\nCultures
(https://t.co/xD07uUAxVj) change. No one would have believe th
at Ireland... https://t.co/gwlxzWe4YY', (https://t.co/gwlxzWe4YY),
'Last week I went to the Norval Foundation and I was drawn to this woode
n sculpture mme Noria Mabasa made after the... https://t.co/VNAEgsIS3r] (h
ttps://t.co/VNAEgsIS3r)]
```

In [ ]:

