

# Introduction to Tweepy

Analysis of Social Media Contents

Alessandro Ortis - University of Catania

## Authentication

Twitter API authentication through the tweepy module. To get your own consumer key, consumer secret, access token and access secret, create a Twitter application: <https://apps.twitter.com/app/new> (<https://apps.twitter.com/app/new>).

```
In [1]: import configparser
        config = configparser.ConfigParser()

        # I put my tokens in a .ini file
        config.read('TwitterConfig.ini')
        auth_info = config['AO_Tutorial']

        CONSUMER_KEY = auth_info['consumer_key']
        CONSUMER_SECRET = auth_info['consumer_secret']
        ACCESS_TOKEN = auth_info['access_token']
        ACCESS_TOKEN_SECRET = auth_info['access_token_secret']

        # You can instead insert them manually
        #CONSUMER_KEY = 'INSERT_YOUR_OWN_HERE'
        #CONSUMER_SECRET = 'INSERT_YOUR_OWN_HERE'
        #ACCESS_TOKEN = 'INSERT_YOUR_OWN_HERE'
        #ACCESS_TOKEN_SECRET = 'INSERT_YOUR_OWN_HERE'
```

```
In [2]: # OAuth is an open standard for access delegation,  
#commonly used as a way for Internet users to grant websites or applications a  
#to their information on other websites but without giving them the passwords.  
  
from tweepy import (Stream, OAuthHandler)  
import tweepy  
  
# Twitter requires all requests to use OAuth for authentication.  
auth = tweepy.OAuthHandler(CONSUMER_KEY, CONSUMER_SECRET)  
auth.set_access_token(ACCESS_TOKEN, ACCESS_TOKEN_SECRET)  
  
api = tweepy.API(auth)  
  
user_account = api.me()  
for key in vars(user_account):  
    print(key)
```

```
_api
_json
id
id_str
name
screen_name
location
profile_location
description
url
entities
protected
followers_count
friends_count
listed_count
created_at
favourites_count
utc_offset
time_zone
geo_enabled
verified
statuses_count
lang
status
contributors_enabled
is_translator
is_translation_enabled
profile_background_color
profile_background_image_url
profile_background_image_url_https
profile_background_tile
profile_image_url
profile_image_url_https
profile_link_color
profile_sidebar_border_color
profile_sidebar_fill_color
profile_text_color
profile_use_background_image
has_extended_profile
default_profile
default_profile_image
following
follow_request_sent
notifications
translator_type
withheld_in_countries
suspended
needs_phone_verification
```

```
In [3]: tweepy.__version__
```

```
Out[3]: '3.10.0'
```

```
In [4]: print("User ID:\t" + str(user_account.id))
print("Name:\t\t" + user_account.name)
print("Screen name:\t" + user_account.screen_name)
print("# of friends:\t"+ str(user_account.friends_count))
print(user_account.profile_image_url)
```

```
User ID:          317046418
Name:            Alessandro
Screen name:     AlessandroOrtis
# of friends:    556
http://pbs.twimg.com/profile_images/791710217197191169/C8Wfn0Az_normal.jpg
```

## User timeline

This example will download the user's home timeline tweets and print each one of their texts.

```
In [5]: public_tweets = api.home_timeline()
        for t_i, tweet in enumerate(public_tweets):
            print("#"+str(t_i)+" "+str(tweet.created_at)+ "\n"+ tweet.user.name + " j
            ust tweeted: " + tweet.text+"\n")
```

#0 2023-02-06 16:30:01

Tuğba Özerk just tweeted: RT @CaglaOzdoker: ÇOK YAKIN DOSTUM HİKMET AĞCA (K. O) ENKAZ ALTINDAYMIŞ HABER ALAMIYORUZ DESTEK OLUN LÜTFEN ULAŞMAMIZA MAHMUT YA NARAY CADDES...

#1 2023-02-06 16:28:00

Farnesina IT just tweeted: La comunità italiana in Brasile costituisce un forte e collante per i rapporti bilaterali ITBR , che si sviluppano a tutto campo.

#2 2023-02-06 16:27:59

Farnesina IT just tweeted: Sds @GiorgioSilli ha ricevuto alla #Farnesina l'Ambasciatore brasiliano Helio Ramos. In Brasile risiedono più di 70... <https://t.co/Hcxz3aiZA2>

#3 2023-02-06 16:27:33

Joe Savino NIGHT TALK with Joe Roxx Thursdays 10PM just tweeted: NOW PLAYING <https://t.co/UeiODupnDv> Gimme All Your Lovin' (Remastered) by @ZZTop DON'T MISS OUR TALK at NIGHT SHOWS <https://t.co/F4kCHbCuAx>

#4 2023-02-06 16:27:19

المحمي مهدي العنزي just tweeted: RT @alkhalid\_news: اللهم احفظه بحفظك kw <https://t.co/2FomseUDAa>

#5 2023-02-06 16:27:03

MONTREAL.AI just tweeted: OPEN EDITION !

Game of Life 🧠 🧠 🧠

[ M I N T I S L I V E ]

Price: 0.0042069 Eth

Mint: <https://t.co/OGwS1olw16>

O... <https://t.co/9GTarqLYos>

#6 2023-02-06 16:25:32

Repubblica just tweeted: Allergica al latte, muore dopo dieci giorni di coma: aveva mangiato un tiramisù "vegano" contaminato da latticini [... <https://t.co/ZEN7Trjggn>

#7 2023-02-06 16:25:13

Repubblica just tweeted: Allergica al latte, muore dopo dieci giorni di coma: aveva mangiato un tiramisù "vegano" contaminato da latticini <https://t.co/QofuhjupHb>

#8 2023-02-06 16:24:33

Joe Savino NIGHT TALK with Joe Roxx Thursdays 10PM just tweeted: NOW PLAYING <https://t.co/UeiODupnDv> - Dreaming of you by @SelenaLaLeyenda DON'T MISS OUR TALK at NIGHT SHOWS <https://t.co/UqCEZLDm5F>

#9 2023-02-06 16:24:28

Joe Savino NIGHT TALK with Joe Roxx Thursdays 10PM just tweeted: LISTEN IN--&gt;&gt; <https://t.co/UeiODupnDv> <https://t.co/01Y5XLNewN>

#10 2023-02-06 16:21:32

Joe Savino NIGHT TALK with Joe Roxx Thursdays 10PM just tweeted: LISTEN IN--&

gt;&gt; <https://t.co/UeiODupnDv> <https://t.co/EQj8kua4xB>

#11 2023-02-06 16:20:47

I Blog del Fatto just tweeted: "La grande gelata che sta investendo l'Italia potrebbe adattarsi anche alla media settimanale delle intenzioni di v... <https://t.co/IBfAVJ5g9j>

#12 2023-02-06 16:18:35

Repubblica just tweeted: Tim down, ecco cosa è successo alla rete e quali possibilità di ottenere un rimborso [di Alessandro Longo] <https://t.co/pt8pa5DnZX>

#13 2023-02-06 16:18:33

Joe Savino NIGHT TALK with Joe Roxx Thursdays 10PM just tweeted: NOW PLAYING <https://t.co/UeiODupnDv> Turn Up The Radio by @AutographBand DON'T MISS OUR TALK at NIGHT SHOWS <https://t.co/1LcWj2RkL9>

#14 2023-02-06 16:18:22

Joe Savino NIGHT TALK with Joe Roxx Thursdays 10PM just tweeted: Autograph - Turn Up The Radio LISTEN IN--&gt;&gt; <https://t.co/UeiODupnDv> VISIT <https://t.co/kjD1cQHudg>

#15 2023-02-06 16:18:10

Il Fatto Quotidiano just tweeted: Il recente attacco è solo l'ultimo episodio: nel 2022 gli hacker hanno sottratto criptovalute per 3,8 miliardi. Di... <https://t.co/lKqeVbdzd8>

#16 2023-02-06 16:17:16

Il Fatto Quotidiano just tweeted: #USA, disoccupazione ai minimi dal 1953: Powell alzando i tassi fa lotta di classe per i "padroni". Di... <https://t.co/BM3CBbu4KP>

#17 2023-02-06 16:16:42

MONTREAL.AI just tweeted: Gm 🤖 🧠 🤖

#18 2023-02-06 16:16:27

Okezone just tweeted: Korban Tewas Gempa Turki dan Suriah Terus Bertambah, 2.310 Tewas Lebih dari 10.000 Luka #LengkapCepatBeritanya... <https://t.co/53jSzULSOQ>

```
In [6]: tweet.__dir__()
```



```
Out[6]: ['_api',
'_json',
'created_at',
'id',
'id_str',
'text',
'truncated',
'entities',
'source',
'source_url',
'in_reply_to_status_id',
'in_reply_to_status_id_str',
'in_reply_to_user_id',
'in_reply_to_user_id_str',
'in_reply_to_screen_name',
'author',
'user',
'geo',
'coordinates',
'place',
'contributors',
'is_quote_status',
'retweet_count',
'favorite_count',
'favorited',
'retweeted',
'possibly_sensitive',
'possibly_sensitive_appealable',
'lang',
'__module__',
'parse',
'destroy',
'retweet',
'retweets',
'favorite',
'__eq__',
'__ne__',
'__doc__',
'__hash__',
'__init__',
'__getstate__',
'parse_list',
'__repr__',
'__dict__',
'__weakref__',
'__str__',
'__getattr__',
'__setattr__',
'__delattr__',
'__lt__',
'__le__',
'__gt__',
'__ge__',
'__new__',
'__reduce_ex__',
'__reduce__',
'__subclasshook__',
```

```
'__init_subclass__',
'__format__',
'__sizeof__',
'__dir__',
'__class__']
```

```
In [7]: tweet.entities
```

```
Out[7]: {'hashtags': [],
'symbols': [],
'user_mentions': [],
"urls": [{'url': 'https://t.co/uDizWk8dSA',
'expanded_url': 'https://twitter.com/i/web/status/1494600424044830745',
'display_url': 'twitter.com/i/web/status/1...',
"indices': [117, 140]}]}
```

The *home\_timeline()* method returns the 20 most recent statuses, including retweets, posted by the authenticating user and that user's friends.

#### Parameters:

**since\_id** – Returns only statuses with an ID greater than (that is, more recent than) the specified ID.

**max\_id** – Returns only statuses with an ID less than (that is, older than) or equal to the specified ID.

**count** – Specifies the number of statuses to retrieve.

**page** – Specifies the page of results to retrieve. Note: there are pagination limits.

**Return type:** list of Status objects

```
In [8]: #for key in vars(tweet):
#       print(key)

print(tweet.id)
print(tweet.user.name)
print(tweet.favorite_count)
print(tweet.retweet_count)
```

```
1622630307932565504
Okezone
0
0
```

## Twitter trends

The method *trends\_closests()* returns the locations that Twitter has trending topic information for, closest to a specified location.

The response is an array of “locations” that encode the location's WOEID and some other human-readable information such as a canonical name and country the location belongs in.

A WOEID is a Yahoo! Where On Earth ID.

```
In [9]: # Rome coordinates
lat = 41.890251
long = 12.492373
location = api.trends_closest(lat, long)

print('# of locations:\t'+str(len(location)))
loc_name = location[0]['name']
loc_id = location[0]['woeid']
print(loc_name)
print(loc_id)
```

```
# of locations: 1
Rome
721943
```

Another way to get WOEID of a location is to run the procedure *trends\_available()*. It returns the locations that Twitter has trending topic information for. The response is an array of “locations” that encode the location’s WOEID (a Yahoo! Where On Earth ID) and some other human-readable information such as a canonical name and country the location belongs in.

```
In [10]: # List of Locations that Twitter has trending topic information for
trend_locations = api.trends_available()
my_place = {}

print('# of places:\t'+str(len(trend_locations)))

for place in trend_locations:
    print(place['name']+" "+str(place["woeid"]))
    if place['name'] == "Rome":
        my_place = place
        break
print('\n')
print(my_place['name'])
print(my_place['woeid'])

print(my_place)
```

# of places: 467  
Worldwide 1  
Winnipeg 2972  
Ottawa 3369  
Quebec 3444  
Montreal 3534  
Toronto 4118  
Edmonton 8676  
Calgary 8775  
Vancouver 9807  
Birmingham 12723  
Blackpool 12903  
Bournemouth 13383  
Brighton 13911  
Bristol 13963  
Cardiff 15127  
Coventry 17044  
Derby 18114  
Edinburgh 19344  
Glasgow 21125  
Hull 25211  
Leeds 26042  
Leicester 26062  
Liverpool 26734  
Manchester 28218  
Middlesbrough 28869  
Newcastle 30079  
Nottingham 30720  
Plymouth 32185  
Portsmouth 32452  
Preston 32566  
Sheffield 34503  
Stoke-on-Trent 36240  
Swansea 36758  
London 44418  
Belfast 44544  
Santo Domingo 76456  
Guatemala City 83123  
Acapulco 110978  
Aguascalientes 111579  
Chihuahua 115958  
Mexico City 116545  
Ciudad Juarez 116556  
Nezahualcóyotl 116564  
Culiacán 117994  
Ecatepec de Morelos 118466  
Guadalajara 124162  
Hermosillo 124785  
León 131068  
Mérida 133327  
Mexicali 133475  
Monterrey 134047  
Morelia 134091  
Naucalpan de Juárez 134395  
Puebla 137612  
Querétaro 138045  
Saltillo 141272

San Luis Potosí 144265  
Tijuana 149361  
Toluca 149769  
Zapopan 151582  
Mendoza 332471  
Santiago 349859  
Concepcion 349860  
Valparaiso 349861  
Bogotá 368148  
Cali 368149  
Medellín 368150  
Barranquilla 368151  
Quito 375732  
Guayaquil 375733  
Caracas 395269  
Maracaibo 395270  
Maracay 395271  
Valencia 395272  
Barcelona 395273  
Ciudad Guayana 395275  
Turmero 395277  
Lima 418440  
Brasília 455819  
Belém 455820  
Belo Horizonte 455821  
Curitiba 455822  
Porto Alegre 455823  
Recife 455824  
Rio de Janeiro 455825  
Salvador 455826  
São Paulo 455827  
Campinas 455828  
Fortaleza 455830  
Goiânia 455831  
Manaus 455833  
São Luís 455834  
Guarulhos 455867  
Córdoba 466861  
Rosario 466862  
Barquisimeto 468382  
Maturín 468384  
Buenos Aires 468739  
Gdańsk 493417  
Kraków 502075  
Lodz 505120  
Poznań 514048  
Warsaw 523920  
Wroclaw 526363  
Vienna 551801  
Cork 560472  
Dublin 560743  
Galway 560912  
Bordeaux 580778  
Lille 608105  
Lyon 609125  
Marseille 610264  
Montpellier 612977

Nantes 613858  
Paris 615702  
Rennes 619163  
Strasbourg 627791  
Toulouse 628886  
Berlin 638242  
Bremen 641142  
Dortmund 645458  
Dresden 645686  
Dusseldorf 646099  
Essen 648820  
Frankfurt 650272  
Hamburg 656958  
Cologne 667931  
Leipzig 671072  
Munich 676757  
Stuttgart 698064  
Bologna 711080  
Genoa 716085  
Milan 718345  
Naples 719258  
Palermo 719846  
Rome 721943

Rome  
721943

```
{'name': 'Rome', 'placeType': {'code': 7, 'name': 'Town'}, 'url': 'http://where.yahooapis.com/v1/place/721943', 'parentid': 23424853, 'country': 'Italy', 'woeid': 721943, 'countryCode': 'IT'}
```

Now we can use the *trends\_place()* procedure that returns the top 50 trending topics for a specific WOEID, if trending information is available for it.

The response is an array of “trend” objects that encode the name of the trending topic, the query parameter that can be used to search for the topic on Twitter Search, and the Twitter Search URL.

This information is cached for 5 minutes. Requesting more frequently than that will not return any more data, and will count against your rate limit usage.

The *tweet\_volume* for the last 24 hours is also returned for many trends if this is available.

```
In [11]: res = api.trends_place(loc_id)
trends = res[0]['trends']
num_trends = len(trends)
print('# of trends:\t'+str(num_trends))

for t_i in range(min(10,num_trends)):
    print(trends[t_i])
```

```
# of trends:    45
{'name': '#amici22', 'url': 'http://twitter.com/search?q=%23amici22', 'promoted_content': None, 'query': '%23amici22', 'tweet_volume': 17668}
{'name': '#sanremo23', 'url': 'http://twitter.com/search?q=%23sanremo23', 'promoted_content': None, 'query': '%23sanremo23', 'tweet_volume': None}
{'name': 'Arisa', 'url': 'http://twitter.com/search?q=Arisa', 'promoted_content': None, 'query': 'Arisa', 'tweet_volume': 12642}
{'name': 'Cricca', 'url': 'http://twitter.com/search?q=Cricca', 'promoted_content': None, 'query': 'Cricca', 'tweet_volume': None}
{'name': 'Paky', 'url': 'http://twitter.com/search?q=Paky', 'promoted_content': None, 'query': 'Paky', 'tweet_volume': None}
{'name': 'Maddalena', 'url': 'http://twitter.com/search?q=Maddalena', 'promoted_content': None, 'query': 'Maddalena', 'tweet_volume': None}
{'name': '#Zelensky', 'url': 'http://twitter.com/search?q=%23Zelensky', 'promoted_content': None, 'query': '%23Zelensky', 'tweet_volume': 13488}
{'name': '#ARMYSelcaDay', 'url': 'http://twitter.com/search?q=%23ARMYSelcaDay', 'promoted_content': None, 'query': '%23ARMYSelcaDay', 'tweet_volume': 12311}
{'name': '#ARSD', 'url': 'http://twitter.com/search?q=%23ARSD', 'promoted_content': None, 'query': '%23ARSD', 'tweet_volume': 12867}
{'name': 'Lavinia', 'url': 'http://twitter.com/search?q=Lavinia', 'promoted_content': None, 'query': 'Lavinia', 'tweet_volume': None}
```

Let's make the content of *trends* more readable.



```
In [12]: for t in trends:
          print(t['name'])
          print("Last 24h tweet volume:\t" +str(t['tweet_volume']))
          print(t['url'])
          print("\n")
```

#amici22  
Last 24h tweet volume: 17668  
<http://twitter.com/search?q=%23amici22>

#sanremo23  
Last 24h tweet volume: None  
<http://twitter.com/search?q=%23sanremo23>

Arisa  
Last 24h tweet volume: 12642  
<http://twitter.com/search?q=Arisa>

Cricca  
Last 24h tweet volume: None  
<http://twitter.com/search?q=Cricca>

Paky  
Last 24h tweet volume: None  
<http://twitter.com/search?q=Paky>

Maddalena  
Last 24h tweet volume: None  
<http://twitter.com/search?q=Maddalena>

#Zelensky  
Last 24h tweet volume: 13488  
<http://twitter.com/search?q=%23Zelensky>

#ARMYSelcaDay  
Last 24h tweet volume: 12311  
<http://twitter.com/search?q=%23ARMYSelcaDay>

#ARSD  
Last 24h tweet volume: 12867  
<http://twitter.com/search?q=%23ARSD>

Lavinia  
Last 24h tweet volume: None  
<http://twitter.com/search?q=Lavinia>

Amadeus  
Last 24h tweet volume: None  
<http://twitter.com/search?q=Amadeus>

Ricky Martin  
Last 24h tweet volume: None

<http://twitter.com/search?q=%22Ricky+Martin%22>

Alessio

Last 24h tweet volume: None

<http://twitter.com/search?q=Alessio>

Amatrice

Last 24h tweet volume: None

<http://twitter.com/search?q=Amatrice>

Gianni

Last 24h tweet volume: None

<http://twitter.com/search?q=Gianni>

Martina

Last 24h tweet volume: 13514

<http://twitter.com/search?q=Martina>

Paola e Chiara

Last 24h tweet volume: None

<http://twitter.com/search?q=%22Paola+e+Chiara%22>

Elodie

Last 24h tweet volume: None

<http://twitter.com/search?q=Elodie>

Isobel

Last 24h tweet volume: None

<http://twitter.com/search?q=Isobel>

La Rai

Last 24h tweet volume: None

<http://twitter.com/search?q=%22La+Rai%22>

Mara Sattai

Last 24h tweet volume: None

<http://twitter.com/search?q=%22Mara+Sattai%22>

Anna Oxa

Last 24h tweet volume: None

<http://twitter.com/search?q=%22Anna+Oxa%22>

Mondiali

Last 24h tweet volume: None

<http://twitter.com/search?q=Mondiali>

Samu

Last 24h tweet volume: None

<http://twitter.com/search?q=Samu>

San Valentino

Last 24h tweet volume: None

<http://twitter.com/search?q=%22San+Valentino%22>

Ariston

Last 24h tweet volume: None

<http://twitter.com/search?q=Ariston>

Premier League

Last 24h tweet volume: 165194

<http://twitter.com/search?q=%22Premier+League%22>

De Zerbi

Last 24h tweet volume: None

<http://twitter.com/search?q=%22De+Zerbi%22>

#Terremoto

Last 24h tweet volume: 97488

<http://twitter.com/search?q=%23Terremoto>

#Terremoto

Last 24h tweet volume: 97488

<http://twitter.com/search?q=%23Terremoto>

#tommasozorzi

Last 24h tweet volume: None

<http://twitter.com/search?q=%23tommasozorzi>

#Brignone

Last 24h tweet volume: None

<http://twitter.com/search?q=%23Brignone>

#oggiueunaltrogiorno

Last 24h tweet volume: None

<http://twitter.com/search?q=%23oggiueunaltrogiorno>

#courchevelmeribel2023

Last 24h tweet volume: None

<http://twitter.com/search?q=%23courchevelmeribel2023>

#6febbraio

Last 24h tweet volume: None  
<http://twitter.com/search?q=%236febbraio>

#ManchesterCity  
Last 24h tweet volume: None  
<http://twitter.com/search?q=%23ManchesterCity>

#lidomengoni  
Last 24h tweet volume: None  
<http://twitter.com/search?q=%23lidomengoni>

#fisalpine  
Last 24h tweet volume: None  
<http://twitter.com/search?q=%23fisalpine>

#r101  
Last 24h tweet volume: None  
<http://twitter.com/search?q=%23r101>

#HogwartsLegacy  
Last 24h tweet volume: 51572  
<http://twitter.com/search?q=%23HogwartsLegacy>

#EurosportSCI  
Last 24h tweet volume: None  
<http://twitter.com/search?q=%23EurosportSCI>

#maremoto  
Last 24h tweet volume: None  
<http://twitter.com/search?q=%23maremoto>

#maneskin  
Last 24h tweet volume: None  
<http://twitter.com/search?q=%23maneskin>

#giuliasalemi  
Last 24h tweet volume: None  
<http://twitter.com/search?q=%23giuliasalemi>

#Tsunami  
Last 24h tweet volume: None  
<http://twitter.com/search?q=%23Tsunami>

## Exercise

Choose a location, find its WOEID, list the recent Twitter trends based on the chosen location and crawl some related Tweets.