

- crear correo con tema Investigación DESI x
- Crear Github vinculado a correo x
- Vincular con Github con tu framework de anaconda (Spyder o VS code o Jupyter) x
- Correr códigos de los dos links de arriba x

<https://github.com/JillieChang/Analisis-de-sentimiento-Python>

[Analisis-de-sentimiento-Python](#) – Github

Usa tweepy

1. Tweepy_Chile.ipynb - Extracción de Tweets en Chile:

Este archivo utiliza **Tweepy**, una librería de Python para interactuar con la API de Twitter, con el fin de extraer tweets relacionados con el COVID-19 desde Chile (el tema de este proyecto).

Pasos importantes en el código:

1. Autenticación:

- Se establece una conexión con la API de Twitter usando claves y tokens, que permiten interactuar con la plataforma y extraer datos.

2. Consulta y extracción:

- Se utiliza una consulta personalizada a la API de Twitter para buscar tweets con ciertos términos relacionados con el COVID-19 (palabras clave como "coronavirus", "cuarentena", etc.) dentro de una ubicación específica, en este caso, Chile.
- El parámetro geocode es clave en la extracción de datos porque permite limitar los resultados a tweets geolocalizados en una región específica (Chile en este caso).
- El script incluye ciclos para extraer grandes volúmenes de tweets, manejando la paginación de la API de Twitter.

3. Almacenamiento de datos:

- Los tweets extraídos se almacenan en un DataFrame de **Pandas**, que incluye columnas como el texto del tweet, la fecha y hora, el número de retweets y favoritos, entre otros.

4. Preprocesamiento:

- El texto de los tweets pasa por un proceso de limpieza para eliminar elementos no deseados, como enlaces o menciones, dejando solo el texto relevante para el análisis posterior.

Resultados personales: No se logró ejecutar el código configurado en el repositorio debido a errores en las librerías usadas, pero se tratará de utilizar tweepy que ya había investigado sobre

ella, aunque tengo entendido se necesita acceso a la API de twitter para usarla, posiblemente sea esa la razón que no funciona.

<https://github.com/Claudio9701/web-scraping-sentiment-analysis/blob/master/Sesi%C3%B3n%206%20-%20Scraping%20de%20Redes%20Sociales%20y%20An%C3%A1lisis%20de%20Sentimiento.ipynb>

[web-scraping-sentiment-analysis](#) – Github

Notebook con un ejercicio para web scraping

1. Importación de librerías:

- **requests:** Para hacer peticiones HTTP y obtener el contenido de páginas web.
- **BeautifulSoup:** Para analizar y extraer datos del HTML.
- **pandas:** Para manejar y almacenar datos en formato tabular.

2. Extracción de datos con requests:

- Se envía una solicitud HTTP a una página web, obteniendo su contenido en formato HTML.
- La función `requests.get(url)` descarga la página.

3. Análisis del contenido HTML:

- El HTML obtenido se pasa a **BeautifulSoup**, que permite buscar elementos dentro del código HTML utilizando selectores como etiquetas, clases o identificadores.
- Se extraen datos específicos, como comentarios o publicaciones de la página web.

Resultados personales: Con el ejemplo dado no se obtiene la fecha de la publicación como se esperaba, por lo que no se extraen los datos web correctamente. Posiblemente a un cambio en facebook.

[\(239\) 3 Ways to Web Scrape Twitter Data and Tweets in 2024 - YouTube](#)

Bardeen.ai

Scraping de un perfil:



table.csv

Name	Handle	Headline	Location	Website	Followers(num)	Followers(link)	Following(num)	Following(link)	Profile image	url
LASA Migraciones, Desplazamientos y Movilidades	@LASAMIGRATION	Promovemos el debate académico sobre las corrientes migratorias que se originan o tienen como destino América Latina y el Caribe @lasacongress		lasaweb.org/es/sections/inâ€¦	1,540	ON/verified_followers	440	ON/following	https://pbs.twimg.com/profile_images/1507393256228892679/m-DaQW1y_200x200.jpg	https://x.com/LASAMIGRATION

Scraping tweets en base a un keyword:



tweetsScrapingBard
eenAI.csv

	A	B	C	D	E	F	G	H	I	J	K	L
1	Name	Handle	Media URL	Retweets	Likes	Comments	Views	Tweet URL	Profile Link	Post Body	Date	Timestamp
2	Iamyesyouareno	@iamyesyouareno	https://pbs.twimg.com/media/GTP2g1HxgAAegTO?format=jpg&name=360x360, https://pbs.twimg.com/media/GTP2g1HxgAAegTO?format=jpg&name=360x360	7K	28K	280	759K	https://x.com/iamyesyouareno/status/1816066123613516101	https://x.com/iamyesyouareno	The UN wrote a 177-page report on replacement migration. Itâ€™s not a conspiracy theory, itâ€™s an open agenda.		2024-07-24T11:01:19.000Z
3	Spryness Migration	@spryness19139	https://pbs.twimg.com/media/GYJVogUX0AAIUvc?format=jpg&name=360x360			1		https://x.com/spryness19139/status/1838125716065176484317	https://x.com/spryness19139	Wake up. Ready to make the Netherlands your home? Hereâ€™s a quick guide to get you started:	24-Jul	2024-07-23T07:58:27.000Z
4	CrystalStar	@ABottleofStars	https://pbs.twimg.com/media/D1zGYKMucAAOR-?format=jpg&name=small, https://pbs.twimg.com/media/D1zGYKMucAAOR-?format=jpg&name=small			6	2	https://x.com/ABottleofStars/status/1106977734163480576	https://x.com/ABottleofStars	Determine Your Purpose: Is it for work, study, or family So thereâ€™s a monarch migration going on and we have an orange tree thatâ€™s blooming! I have to go to work in 10minutes but here are some quick snaps I got before running out the door	23-Sep	2019-03-16T17:57:13.000Z
5	Daily Romania	@daily_romanian	https://pbs.twimg.com/media/GSOTSDWXYAAkFA0?format=jpg&name=small, https://pbs.twimg.com/media/D1zo7QUgAA2_15?format=jpg&name=small,	905	10K			https://x.com/daily_romanian/status/1811453416041812238	https://x.com/daily_romanian	I predict mass migration of Western Europeans to Eastern Europe in the coming decades.	16-Mar-19	2024-07-11T17:32:04.000Z
6	kimmy	@yummykimmy	https://pbs.twimg.com/media/D1zo7QUgAA2_15?format=jpg&name=small			3		https://x.com/yummykimmy/status/1107015719932186624	https://x.com/yummykimmy	Do people not understand at least the concept of migration	11-Jul	2019-03-16T20:28:09.000Z
7												



tweet Scapping
Bardeen2.csv

	A	B	C	D	E	F	G	H	I	J	K	L
1	Name	Handle	Media URL	Retweets	Likes	Comment	Views	Tweet URL	Profile Link	Post Body	Date	Timestamp
1	Dan Gamboa Boh's z...	@DanGamboaB	https://pbs.twimg.com/media/EZ6-owWXgAly5T?format=jpg&name=360x360, https://pbs.twimg.com/media/EZ6-owWXgAly5T?format=jpg&name=360x360	5.5K	29K	685		https://x.com/DanGamboaB/status/1269661783477886978	https://x.com/DanGamboaB	Sabias que durante la Pangea, Mxico y Colombia estaban unidos y esto explica porqu Colombia tiene mariachis? La gran migracin de mariachis sucedi hace 175 millones de aos durante la Era Paleozoica. (1/2) Firm un convenio de cooperacin con la @OIM_Mexico para reforzar acciones que benefician a las empresas locales y promuevan un entorno Ms inclusivo que fortalezca nuestro compromiso con la comunidad migrante y contribuya al bienestar de todos.	#####	2020-06-07T16:05:33.000Z
2	Octavio de la Torre	@octaviodlortor	https://pbs.twimg.com/media/GYQho2tXQA A24qm?format=jpg&name=360x360, https://pbs.twimg.com/media/GYQho2tXQA A24qm?format=jpg&name=360x360		2		1	https://x.com/octaviodlortor/status/1838631536184242562	https://x.com/octaviodlortor	#CONCANACO #Migracin Plutarco Elas Calles fu un "sefard" judeo-masn grado 33 que fund el Partido Revolucionario Institucional (PRI) https://sdnnoticias.com/columnas/presidente-calles-calles.html https://nacion321.com/politica-1/quienes-son-los-politicos-mexicanos-masonesC #Europa #pennsula #iberica #judos #sefard-s #marranos #criptojudos #migracin #America #Latina #Mxico #hispanos	37m	2024-09-24T17:28:13.000Z
3	megafenix64	@megafenix64	https://pbs.twimg.com/media/GYHO--zb0AIGpOF?format=jpg&name=small, https://pbs.twimg.com/media/GYHO--zb0AIGpOF?format=jpg&name=small	13	19		5 1.6K	https://x.com/megafenix64/status/1837979924126486637	https://x.com/megafenix64	CMxico ha dejado de ser un pas de paso para convertirse en destino de miles de migrantes. Es urgente definir polticas de integracin que respeten los derechos de los migrantes y fomenten el empleo formal en igualdad.C @RosyUrbinaTap Desgraciadamente este gobierno de la 4T lo hace todo al revs.		2024-09-22T22:18:56.000Z
4	Cristian Alegria	@CristianAlegria	https://pbs.twimg.com/media/GYBEIFKWwA AgnGf?format=jpg&name=small, https://pbs.twimg.com/media/GYBEIFKWwA AgnGf?format=jpg&name=small				26	https://x.com/CristianAlegria/status/1837544105536020779	https://x.com/CristianAlegria	1- Hacen un programa de ayuda a los migrantes, con la tasa tan grande de desempleo pudiendo hacer programas de empleo para los mexicanos. 2. La cuarta Despus de la pandemia el nmero de migrantes no documentados en Mxico explot.		2024-09-21T17:27:09.000Z
5	Angie Muoz	@angiemunoz79	https://pbs.twimg.com/media/GJSAJEnWEA Ehm40?format=jpg&name=360x360, https://pbs.twimg.com/media/GJSAJEnWEA Ehm40?format=jpg&name=360x360	296	434	44	13K	https://x.com/angiemunoz79/status/1771181455701430745	https://x.com/angiemunoz79	As no T a Licha @aliciabarcena @SRE_mx		2024-03-22T14:25:40.000Z
6	Poltica Bsica	@PoliticaBasica	https://pbs.twimg.com/media/GOw8jWakA AcGDc?format=png&name=small, https://pbs.twimg.com/media/GOw8jWakA AcGDc?format=png&name=small	68	529	37	31K	https://x.com/PoliticaBasica/status/1795483740669702407	https://x.com/PoliticaBasica	1- Hacen un programa de ayuda a los migrantes, con la tasa tan grande de desempleo pudiendo hacer programas de empleo para los mexicanos. 2. La cuarta Despus de la pandemia el nmero de migrantes no documentados en Mxico explot.		2024-05-28T15:54:16.000Z
7										Antes Mxico registraba en promedio 140 mil al ao... En 2023 registr casi 800 mil y se espera que en 2024 se supere el milln. Como pas de paso (y ya tambin destino)C Qu vamos a hacer para	28-May	2024-05-28T15:54:16.000Z

Resultados Personales: Funciona y puede servir para el proyecto, con la configuraci3n para obtener tweets en base a un keyword y guardarlo en un Google sheets, sÃ­ encontr3 algunos tweets, en este cas3 probado con keywords "migraci3n" y "M3xico", tendrÃ­a que ver como se puede utilizar de mejor forma la plataforma, para obtener mÃ¡s registros de publicaciones.

[\(243\) How I get Tweet data for FREE in 2024 as a data scientist - YouTube](#)

```
from twikit import Client, TooManyRequests
import time
from datetime import datetime
import csv
import pathlib
from configparser import ConfigParser
import configparser
from random import randint

MINIMUM_TWEETS = 10
QUERY = 'chatgpt'

## Login credentials
config_path = pathlib.Path(__file__).parent.absolute() / "config.ini"
config = configparser.ConfigParser()
config.read(config_path)
print(config.sections())
print(config['X'])
username = config['X']['username']
email = config['X']['email']
password = config['X']['password']

## authenticate to X.com
#! 1) use the login credentials. 2) use cookies.
client = Client(language='en-US')
client.login(auth_info_1=username, auth_info_2=email, password=password)
client.save_cookies('cookies.json')
```

Resultados Personales: Parece que funcionaba cuando salió el video, pero Twitter lo ha restringido. Al querer guardar las cookies para no ser baneado, da el siguiente error que es comentado por otras personas en el video:

```
raise BadRequest(message, headers=response.headers)
twikit.errors.BadRequest: status: 400, message: '{"errors":[{"code":366,"message":"flow name LoginFlow is currently not accessible"}]}'
```

```
twikit.errors.BadRequest: status: 400, message: '{"errors":[{"code":366,"message":"flow name LoginFlow is currently not accessible"}]}'
```

Investigando, es un error relacionado con que twitter te limita para loggear, cuando entré a la aplicación pensó que era un bot y me hizo verificar. Igualmente Twikit parece ser buena opción vigente que funcione.

Otros repositorios:

<https://github.com/sedpeep/Twitter-Scraper/blob/main/main.py>

<https://github.com/imperatrone/twitter-scraper>

alternativas:

Plataforma: Apify, plataforma de servicio para usar aplicaciones que hacen web scraping. Es posible conseguir todos los datos que sean necesario con el plan gratuito, dependiendo de los datos y la cantidad