

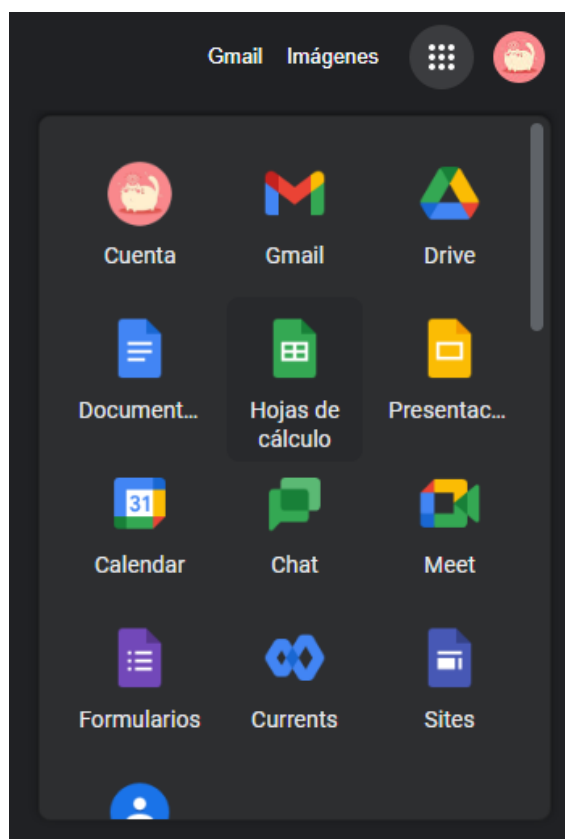
TWEET ARCHIVER:

Recopilación de Datos de la red Social Twitter.

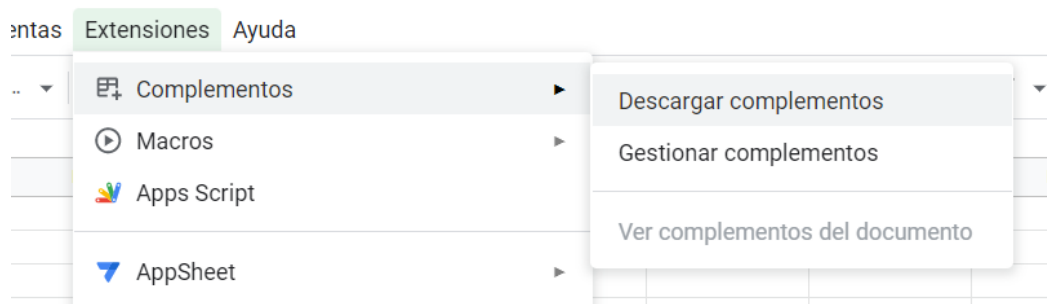
Luis Felipe Narvaez Gomez. E-mail: luis.narvaez@usantoto.edu.co. Cod: 2312660. Facultad de Ingeniería de Sistemas.

GUÍA DE USO BÁSICO - TWEET ARCHIVER

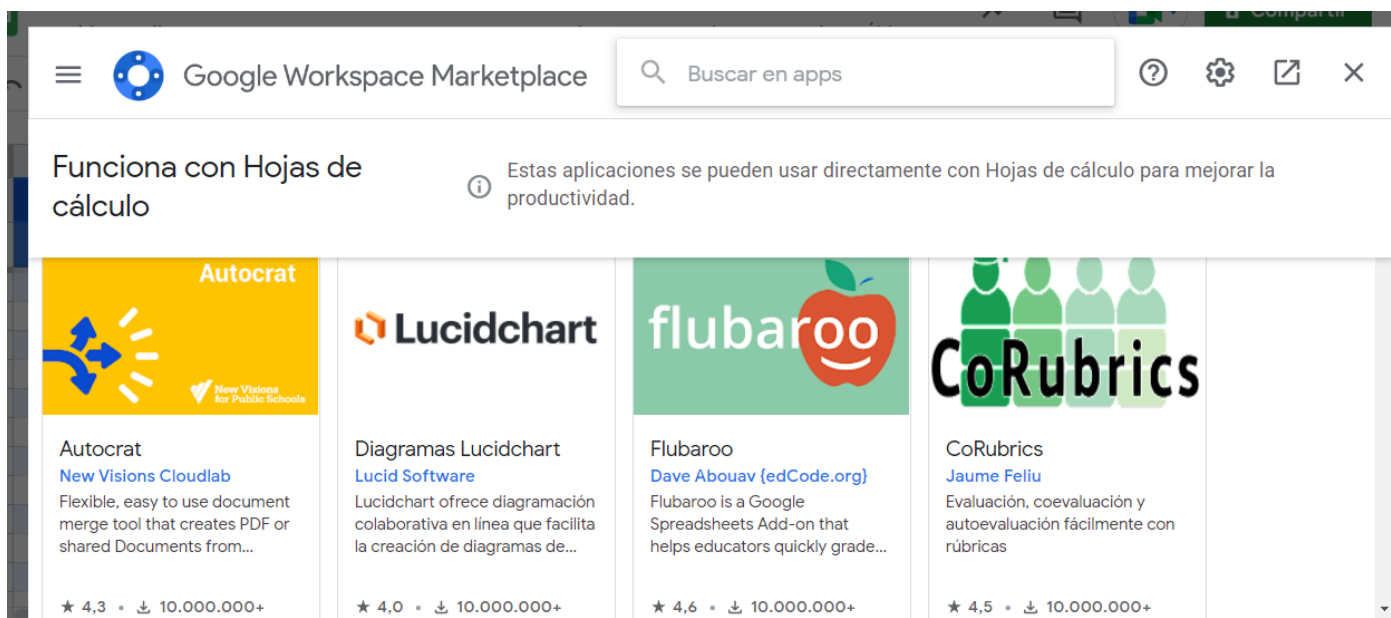
Las hojas de cálculo de las cuales disponemos en Google al crearnos una cuenta de GMAIL poseen varios tipos de exenciones especiales que nos pueden servir para la extracción, recopilación y clasificación de data en bruto de sitios web de interés,, una de estas herramientas es TWEET ARCHIVER la cual nos permite extraer comentarios publicados en la red social de twitter sobre determinados temas específicos, temas en tendencia y hashtags importantes que nosotros necesitemos, a manera de Query que se ira armando según las especificaciones que nosotros le introduzcamos. Para acceder a esta herramienta primero debemos loguearnos en GMAIL, donde se recomienda estar en el navegador nativo de Google, Google Chrome. En la parte superior derecha del navegador encontraremos el menú de despliegue de herramientas, en el encontraremos las hojas de cálculo fácilmente localizable por su icono de color verde, este lo seleccionaremos.



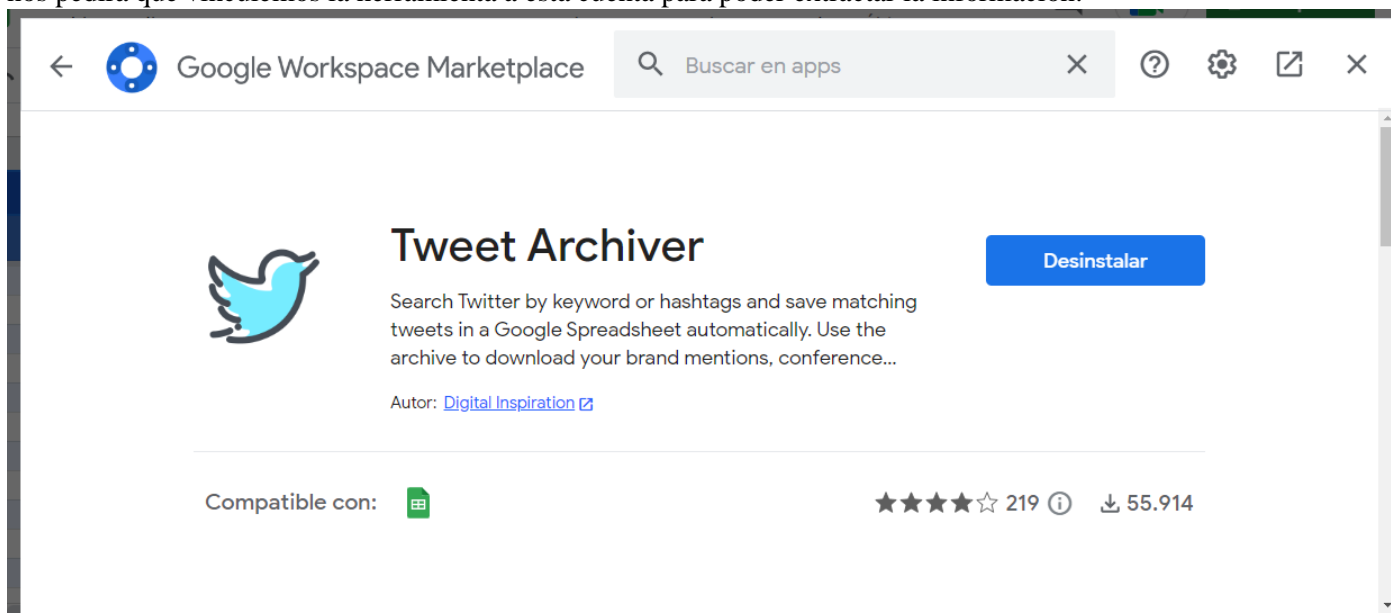
Una vez dentro de la hoja de cálculo nos dirigiremos a la barra de menú en el apartado de Extensiones, en el seleccionamos la herramienta de Complementos y dentro de ella optamos por descargar complementos.



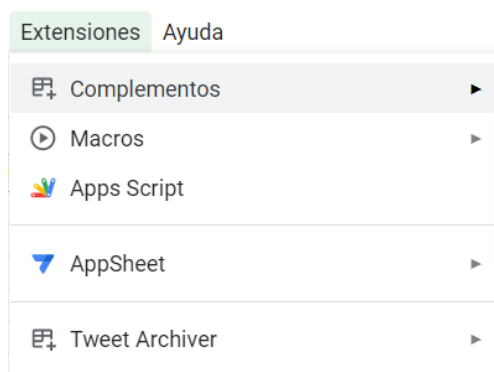
Una vez seleccionamos esta herramienta se nos desplegara una ventana emergente donde podremos buscar la herramienta que necesitamos, en nuestro caso TWEET ARCHIVER.



En la barra de búsqueda buscamos la herramienta que necesitamos por su nombre directo, al encontrarla solo la descargamos. Debemos tener en cuenta que la herramienta se ligará a la hoja de calculo de la cuenta de GMAIL del que estemos logeados, a su vez se recomienda iniciar sesión en TWITTER en el mismo navegador pues a continuación se nos pedirá que vinculemos la herramienta a esta cuenta para poder extraectar la información.



Ahora para poder utilizarla solo vamos al mismo menú de Extensiones y la encontraremos instalada, solo queda abrirla.



Paralelamente en nuestra cuenta de Twitter podemos buscar alguna tendencia o tema de nuestro interés del cual querramos rescatar los datos de las publicaciones en mi caso seleccione el siguiente

Tendencia en Colombia #NosEstanMatando 2.259 Tweets	...
Tendencia en Colombia ESMAD 9.015 Tweets	...

Ahora bien, dentro de nuestra hoja de calculo y con la herramienta de TWEET ARCHIVER abierta, daremos en la opción de crear una nueva regla “créate rule”.

Allí daremos las indicaciones de que tan especifica queremos nuestra query para la búsqueda en la red social. Mientras más específica sea la consulta, como en toda base de datos, menores serán los resultados obtenidos, pero estos serán más de acuerdo a lo que necesitamos, esto podemos observarlo en la configuración de la consulta que hicimos. Pues a pesar de que el hashtag tenga 2.259 Tweets y la palabra ESMAD tenga otros 9.015, realmente rescataremos muchísimos menos tweets propios de la especificidad de la consulta.

Create Twitter Rule

×

All of these words

Esmad

This exact phrase

Any of these words

Esmad, Policia, Colombi

None of these words

These #hashtags

#NosEstanMatando

Written in

Spanish (español)

Near This Place

Premium feature

Advanced Rules

-filter:retweets -filter:repl

People

To these accounts

Mentioning accounts

From these accounts

Twitter Search Query:

Esmad Esmad, OR Policia, OR Colombia #NosEstanMatando -filter:retweets -filter:replies

Create Search Rule

Upgrade to Premium

Cancel

Una vez puestos los datos podemos dar en el botón azul que tal como especifica, creara la regla de búsqueda o el Query y empezara la búsqueda en la red social a través de nuestra cuenta personal. Podemos observar la consulta en la parte inferior, antes de la sección de botones como “Twitter Search Query” la cual al iniciar la búsqueda se reemplazara con un mensaje de “Working, please wait”.

Working, please wait.

Create Search Rule

Upgrade to Premium

Cancel

Luego de esto, obtendremos en nuestra hoja de calculo los datos de los tweets que hayamos rescatado. Sin embargo, aun tendremos unas ventanas emergentes de nuestra herramienta que nos recordaran que la misma tiene una opción de uso gratuito y su capa de pago. Así como el mostrarnos que reglas tenemos creadas hasta el momento, en caso de modificar o crear una nueva.

Create Twitter Rule

You are all set!

The Twitter Archiver is now running and it will fetch tweets once per hour.

You can close this Google sheet and the add-on will continue to run in the background.

Would you like to upgrade?

The premium edition of Twitter Archiver is recommended for power users.

1. You can monitor multiple Twitter search terms and hashtags.
2. It is 4x faster and fetches new tweets every 15 minutes.
3. Need help? You are entitled to complimentary technical support for 60 days.

Should you have any questions, please email support@digitalinspiration.com.

Yes, upgrade to premium

No, maybe later

Manage Twitter Rules

Esmad Esmad, OR Policia, OR Colombia #NosEstanMatando lang:es -filter:retweets -filter:replies

Twitter Archiver is now running in the background and will automatically save new tweets once per hour. Go premium to download new tweets every 15 minutes.

Go Premium!

Edit

Delete

En nuestra hoja de calculo tendremos algo similar a lo siguiente:

Hoja de cálculo sin título										
Archivo Editar Ver Insertar Formato Datos Herramientas Extensiones Ayuda Última modificación ayer a la(s) 23:13										
100% 123 Droid Sans 10 B I A										
A1	Twitter Query: Esmad Esmad, OR Policia, OR Colombia #NosEstanMatando lang:es -filter:retweets -filter:replies									
2	Date	Screen Name	Full Name	Tweet Text	Tweet ID	Link(s)	Media	Location	Retweets	Favorites
3	30/08/2022 02:18	@Retuiteadora3	Retuiteo lo importante!	#Nosestánmatando Petro y su gobierno no dan garantías y tie	1564497877556158465				1	1
4	25/08/2022 19:38	@jelo	jfk	La POLA REGRESO!! Invitado especial @HectorNaupari Hablar	1562857074962608135		https://pbs.twimg.com/media/FbHhLb7Xg		0	0
5	24/08/2022 23:56	@jeshus_77	Jeshús	No sé confundió, el sabe que estamos bajo el gobierno de Juar	1562650016535101440		https://pbs.twimg.com/media/Fa-IS-q50A		1	1
6	24/08/2022 23:51	@jeshus_77	Jeshús	En Colombia #NosEstanMatando Nueva masacre en Buga Vall	1562648798144344066		https://pbs.twimg.com/media/Fa-krMAXw		0	1
7	24/08/2022 23:17	@jeshus_77	Jeshús	#NosEstanMatando Nueva masacre en Colombia, tres person	1562640218993602566	https://twitter.com/Leo_unibist	https://twitter.com/Leo_unibist		2	5
8	23/08/2022 15:32	@misioneroalbert	Carlos Pineda Castro	Asi como fue creada debe ser eliminada la figura del ESMAD e	1562160896507527170	https://twitter.com/Leo_u...			19	18
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										
28										

Solo queda descargar el archivo como un archivo de Excel, no directamente como CSV pues aun necesitamos limpiar la data que tenemos. Cuando nos referimos a limpiar la data hablamos de eliminar todos aquellos caracteres especiales que no necesitamos o pueden presentarse un conflicto al momento de subirlos en la base de datos por su codificación. A su vez también hablamos de eliminar datos que por norma y/o especificación de lo que necesitamos vamos a eliminar las columnas.

Así luce la data original:

Twitter_Especifico_ESMAD (Original Data).xlsx - Excel																
Luis Felipe Narváez Gómez																
No sé confundió, el sabe que estamos bajo el gobierno de Juan Manuel Santos II y que él da órdenes. Pero no olvidamos que #NosEstanMatando hoy hubo otra masacre. Ya van 15 en gobierno de Petro. Y la gente hablando de #SpamDeTenis y Esmad																
Date	Screen Name	Full Name	Tweet Text	Tweet ID	Link(s)	Media	Location	Retweets	Favorites	App	Followers	Follows	Listed	Verified	User Since	Location
30/08/2022 02:18	@Defuileadora3	Rafael	#Nosestammatando Petro y su gobierno no dan garan	1564497877556158465				1	0	1 Twitter for iPhone	3175	3411	0		11/10/2019	
25/08/2022 13:38	@Jefe	Jefu	La POLA REGRESO!! Invitado especial @HectorNaupe	1562857074962608135	https://pbs.twimg.com/media/F8R6LJ			0	0	Twitter Web App	186	373	1		20/11/2009	
24/08/2022 23:56	@Jesusa_17	Jesús	[No sé confundió, el sabe que estamos bajo el gobiern	156265010535101440	https://pbs.twimg.com/media/Fa-E-d			1	1	Twitter for Andro	1112	719	2		4/03/2018	Antioquia, Colombiá SN DOLO PO
24/08/2022 23:51	@Jesusa_17	Jesús	En Colombia #NosEstanMatando Nueva masacre en B	1562644798144344066	https://pbs.twimg.com/media/Fa-JaH			0	1	Twitter for Andro	1112	719	2		4/03/2018	Antioquia, Colombiá SN DOLO PO
24/08/2022 23:17	@Jesusa_17	Jesús	#NosEstanMatando Nueva masacre en Colombia, tres	1562640213993020566	https://twitter.com/Leo_u1/pic.twitter.com/7a6toH0D			2	5	Twitter for Andro	1112	719	2		4/03/2018	Antioquia, Colombiá SN DOLO PO
23/08/2022 15:32	@missperezabell	Carlos Pineda Castro	Así como fue creada debe ser eliminada la figura del E	1562160896507527170	https://pbs.twimg.com/media/Fa3eD			19	18	Twitter for Andro	743	1707	0		26/05/2012	Animador soc

Ahora así luce la data en limpio.

Twitter_Especifico_ESMAD (Data Limpia).xlsx - Excel																
Luis Felipe Narváez Gómez																
date	tweet_text	tweets	favorites	app	location											
30/08/2022 02:18	#Nosestammatando Petro y su gobierno no dan garan	1	0	1 http://twitter.com/download/iphone?Twitter for iPhone												
25/08/2022 13:38	La POLA REGRESO!! Invitado especial @HectorNaupe	0	0	https://mobile.twitter.com?Twitter Web App												
24/08/2022 23:56	No sé confundió el sabe que estamos bajo el gobiern	1	1	1 http://twitter.com/	Antioquia Colombia											
24/08/2022 23:51	En Colombia #NosEstanMatando Nueva masacre en E	0	1	1 http://twitter.com/	Antioquia Colombia											
24/08/2022 23:17	#NosEstanMatando Nueva masacre en Colombia tres	2	5	5 http://twitter.com/	Antioquia Colombia											
23/08/2022 15:32	Así como fue creada debe ser eliminada la figura del E	19	18	18 http://twitter.com/download/android?Twitter for Android												

Una vez tenemos los datos limpios, podemos guardar la hoja como un CSV delimitado por comas, el mismo lo abriremos con un editor de texto básico como Bloc de Notas, esto con la finalidad de corroborar que cada dato de las filas este separado por comas y no por otro carácter, si este ultimo es el caso podemos reemplazar el mismo por la coma.

1	date,tweet_text,tweets,favorites,app,location
2	30/08/2022 02:18,#Nosestammatando Petro y su gobierno no dan garantías y tienen entumecidos al ejército a los policías y al ESMAD,1,1,"http://twitter.com/download/iphone?Twitter for iPhone",
3	25/08/2022 13:38,La POLA REGRESO!! Invitado especial @HectorNaupe! Hablaremos de la incidencia de la elección de pretero en la región los esperamos #ElPeorPresidenteDeLaHistoria
4	#NosEstanMatando #EsJusticia #ESMAD #FelizJueves #libertarios,0,0,"https://mobile.twitter.com?Twitter Web App",
5	24/08/2022 23:56,No se confundió el sabe que estamos bajo el gobierno de Juan Manuel Santos II y que el da órdenes. Pero no olvidamos que #NosEstanMatando hoy hubo otra masacre. Ya van 15 en gobierno de Petro. Y la gente hablando de #SpamDeTenis y Esmad,1,1,"http://twitter.com/download/android?Twitter for Android",Antioquia Colombia
6	24/08/2022 23:51,En Colombia #NosEstanMatando Nueva masacre en Buga Valle la noche de hoy #NosEstanExpropiando Ya le arrebataron tierras al sr alvaro Uribe Velez. ¿No que no iba hacer expropiación?. #PareceUnChiste #SpamDeTenis ESMAD,0,1,"http://twitter.com/download/android?Twitter for Android",Antioquia Colombia
7	24/08/2022 23:17,#NosEstanMatando Nueva masacre en Colombia tres personas fueron asesinadas en Buga Valle la noche de hoy. #505Colombia @ONUHumanRights @ONUderechos @JuanitaGoe #PetroBastaYa #PetroPareLaMasacre #SpamDeTenis ESMAD RT @lco uribista: ??URGENTE: Masacre #11 del Gobierno Petro asesinan 3 personas en Buga...quien dio la orden???,2,5,"http://twitter.com/download/android?Twitter for Android",Antioquia Colombia
8	23/08/2022 15:32,Así como fue creada debe ser eliminada la figura del ESMAD en la policía. No tiene coherencia que @petrogustavo ordene a la fuerza pública proteger la VIDA si se mantiene este brazo letal. #NosEstanMatando cobra mas vigencia que nunca. Prohibido olvidar.@MJDuzan @PizarroMaríaJo,19,18,"http://twitter.com/download/android?Twitter for Android"

Una vez tenemos nuestra data visitamos el sitio web <https://csvjson.com/csv2json> para convertir nuestro archivo CSV a JSON.

Flatfile
csvjson

Turnkey CSV Importer • Installs in minutes

Home / csv2json ▾

CSV or TSV ▾ JSON

To get started, upload or paste your data from Excel (saved as CSV or TSV).

Upload a CSV file

Seleccionar archivo

Ninguno archivo selec.

Or paste your CSV here

```
date,tweet_text,tweets,favorites,app,location
30/08/2022 02:18,#Nosestanmatando Petro y su
gobierno no dan garantias y tienen entumecidos al
ejercito a los policias y al
ESMAD!,1,1,"http://twitter.com/download/iphone"Twit
ter for iPhone",
25/08/2022 13:38,La POLA REGRESO!! Invitado especial
@HectorNaupari Hablaremos de la incidencia de la
eleccion de preto en la region los esperamos
#ElPeorPresidenteDeLaHistoria #NosEstanMatando
#EsJusticia #ESMAD #FelizJueves
#libertarios,0,0,"https://mobile.twitter.com"Twitte
r Web App",
24/08/2022 23:56,No se confundio el sabe que estamos
bajo el gobierno de Juan Manuel Santos ll y que el
da ordenes. Pero no olvidamos que #NosEstanMatando
hoy hubo otra masacre. Ya van 15 en gobierno de
Petro. Y la gente hablando de #SpamDeTenis y
```

Separator

Comma ▾

☒ Parse numbers

☒ Parse JSON

☐ Transpose

Output: ☒ Array ☐ Hash ☐ Minify

JSON

```
[
{
  "date": "30/08/2022 02:18",
  "tweet_text": "#Nosestanmatando Petro y su gobierno no dan garantias y
tienen entumecidos al ejercito a los policias y al ESMAD!",
  "tweets": 1,
  "favorites": 1,
  "app": "http://twitter.com/download/iphone\"Twitter for iPhone",
  "location": ""
},
{
  "date": "25/08/2022 13:38",
  "tweet_text": "La POLA REGRESO!! Invitado especial @HectorNaupari
Hablaremos de la incidencia de la eleccion de preto en la region los
esperamos #ElPeorPresidenteDeLaHistoria #NosEstanMatando #EsJusticia #ESMAD
#FelizJueves #libertarios",
  "tweets": 0,
  "favorites": 0,

```

➤ Convert

📄 Clear

📄 Download


💾 Save

📄 Copy to clipboard

```

1 {
2   "date": "30/08/2022 02:18",
3   "tweet_text": "#NosistanMatando Petro y su gobierno no dan garantias y tienen entumecidos al ejercito a los policias y al ESMAOI",
4   "tweets": 1,
5   "favorites": 1,
6   "app": "http://twitter.com/download/iphone/Twitter for iPhone",
7   "location": ""
8 }
9
10 {
11   "date": "25/08/2022 13:38",
12   "tweet_text": "La POLA REGRESO!! Invitado especial @HectorNaupari Hablaremos de la incidencia de la eleccion de petro en la region los esperamos #ElPeorPresidenteDeLaHistoria #NosistanMat",
13   "tweets": 0,
14   "favorites": 0,
15   "app": "https://mobile.twitter.com/Twitter Web App",
16   "location": ""
17 }
18
19 {
20   "date": "24/08/2022 23:56",
21   "tweet_text": "No se confundo el sabe que estamos bajo el gobierno de Juan Manuel Santos ll y que el da ordenes. Pero no olvidamos que #NosistanMatando hoy hubo otra masacre. Ya van 15 en",
22   "tweets": 1,
23   "favorites": 1,
24   "app": "http://twitter.com/download/android/Twitter for Android",
25   "location": "Antioquia Colombia"
26 }
27
28 {
29   "date": "24/08/2022 23:51",
30   "tweet_text": "En Colombia #NosistanMatando Nueva masacre en Buga Valle la noche de hoy #NosistanExpropiando Ya le arrebaron tierras al sr alvaro Uribe Velez. ❖No que no iba hacer exprop",
31   "tweets": 2,
32   "favorites": 5,
33   "app": "http://twitter.com/download/android/Twitter for Android",
34   "location": "Antioquia Colombia"
35 }
36
37 {
38   "date": "24/08/2022 23:17",
39   "tweet_text": "#NosistanMatando Nueva masacre en Colombia tres personas fueron asesinadas en Buga Valle la noche de hoy. #SOSColombia @ONUHumanRights @ONU_derechos @JuanitaGoe #PetroBastaY",
40   "tweets": 2,
41   "favorites": 5,
42   "app": "http://twitter.com/download/android/Twitter for Android",
43   "location": "Antioquia Colombia"
44 }
45
46 {
47   "date": "23/08/2022 15:32",
48   "tweet_text": "Asi como fue creada debe ser eliminada la figura del ESMAOI en la policia. No tiene coherencia que @petrogustavo ordene a la fuerza publica proteger la VIDA si se mantiene est",
49   "tweets": 19,
50   "favorites": 18,
51   "app": "http://twitter.com/download/android/Twitter for Android",
52   "location": ""
53 }
54 }

```



The screenshot shows a MongoDB command prompt window with the following content:

```

Welcome x db.tweet_esmad.insert() [ ... x
Twitter localhost:27017 test
db.tweet_esmad.insert(
[
  {
    "date": "30/08/2022 02:18",
    "tweet_text": "#Nosestanmatando Petro y su gobierno no dan garantias y tienen entumecidos al ejercito a los policias y al ESMAD!",
    "retweets": 1,
    "favorites": 1,
    "app": "http://twitter.com/download/iphone\"Twitter for iPhone",
    "location": ""
  },
  {
    "date": "25/08/2022 13:38",
    "tweet_text": "La POLA REGRESO!! Invitado especial @HectorNaupari Hablaremos de la incidencia de la eleccion de pretro en la regio
    "retweets": 0,
    "favorites": 0,
    "app": "https://mobile.twitter.com\"Twitter Web App",
    "location": ""
  },
]
0.249 sec.
Inserted 1 record(s) in 248ms

```

Download here.

Welcome x db.tweet_esmad.find() x

Twitter localhost:27017 test

```
db.tweet_esmad.find(
)
```

tweet_esmad 0.006 sec.

```
/* 1 */
{
  "_id" : ObjectId("630ee4bba4409da66684a53e"),
  "date" : "30/08/2022 02:18",
  "tweet_text" : "#Nosestammatando Petro y su gobierno no dan garantias y tienen entumecidos al ejercito a los policias y al ESMAD!",
  "retweets" : 1.0,
  "favorites" : 1.0,
  "app" : "http://twitter.com/download/iphone\"Twitter for iPhone",
  "location" : ""
}

/* 2 */
{
  "_id" : ObjectId("630ee4bba4409da66684a53f"),
  "date" : "25/08/2022 13:38",
  "tweet_text" : "La POLA REGRESO!! Invitado especial @HectorNaupari Hablaremos de la incidencia de la eleccion de pretro en la regio",
  "retweets" : 0.0,
  "favorites" : 0.0,
  "app" : "https://mobile.twitter.com\"Twitter Web App",
  "location" : ""
}

/* 3 */
{
  "_id" : ObjectId("630ee4bba4409da66684a540"),
  "date" : "24/08/2022 23:56",
  "tweet_text" : "No se confundio el sabe que estamos bajo el gobierno de Juan Manuel Santos ll y que el da ordenes. Pero no olvidamo",
  "retweets" : 1.0,

```

< >

Welcome x db.tweet_esmad.count() x

Twitter localhost:27017 test

```
db.tweet_esmad.count(
)
```

0.003 sec.

6