Data Exports: Twitter

Page

Page-level metrics will give you data aggregated at a daily level, so you can look back and see what trends are occurring on your Page over time.

Tweet

Tweet-level metrics show the scope and strength of influence of a tweet, which gives you the ability to analyse the content you're delivering and how well it is resonating with not only your community, but also the potential audience accessible through your community.

Hashtag

Hashtag-level metrics indicate the amount of attention and engagement a certain topic/event is able to evoke.

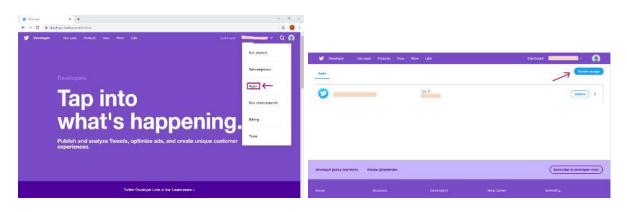
Media

Video-level metrics helps you understand how media contents contribute to your overall reach and engagement.

How do I extract tweets and comments?

Twitter comments, as well as other Twitter data, can be extracted by making use of a Twitter developer's account and R.

First, apply for a Twitter developer account at <u>https://developer.twitter.com</u> and complete the required steps in the application. Once approved, visit your developers page and create an app.



		💓 Developer Motozofa Produkta Ovia Mires Latis Davidicana 🗰 -
Inderstanding appr	App details	Lating and Annual Contract States
n n en app'	*** The start of a maximum of the field of the field of the start of a start of the start of the generation that a 27 bits a start of the start of the start of the start of the start.	https://www.wareautoreautory.
e registes an app?	Any news (second) (0)	-+ had perifier
	This can it was a 22	Terms of Service URL (1)
di products require an API lay?	Application description (working	reput-
	Denote a flow getter of prior agin. This description will be which to carrie as the least quark (Alexantic RC Tenescolet space age does.	Privacy policy URL (1)
	There is shared	republic to the second s
	Take Distance	
		Organization name 🗇
		Organization vebatte URL
		Tell us how this app will be used (required)
	Francis David PB, Association	This least scrap vestice to Scotte any larger any target scattering from the scotter any scattering of the
	Website UBL (resent) (0)	samt What will be existing one and prove contention that the
	228e%	Proceeding School Schoo
	Allow this application to be used to sign in with Twitter (sees rule)	
	🖹 Enoble Sign in with Twitter	
	Calibace URLa	
	OF all the upplications of sub-specify had one of collows (PL on the expect spec	

Now you have a new app! You can always find it on the apps page (developer.twitter.com/en/apps) and click on the "Details" icon to view.

Finally, collect the Application Programming Interface (API) keys under Keys and tokens. Generate your access token and access token secret. All the four keys on this page will be used, so keep them somewhere.

ai 10			
Appricipit	Repuest Indens Pr	alulan	
	ilanan yalar access toilan and i		and receased attention rates before non-sufficient
104.0 10 203	country of party of them at	the state of the second second second	ALC DOMAS AND AND ALL LAS COMPACT SHITTLESS
	Keys and takens		
	last-only period area	Contraction and Contraction of Contr	
	Consumer API kays		(mem)
	API bays	-	
	API secret kay:		-
	Assess falses it assess to	anisonal	C and Co
		energi menyarkan pender yana kanangan kalangan kanangan kananga Kanangan kanangan kana	and the second s
			6
	Access tokes:		10 percent = 2.52
	Access in some sacreet.	Engentinte	

Build a connection between R and Twitter

Install R and R studio first on your device. Both are free!

Open R studio. Now, you want to make use of your Twitter developer's status through R commands, so you give R access to your developer's account.

Tentral Ave.	101	Coherent House	(initialized		
		- Di le mor late			
		Characterist :	111 B		18
Constant C.G. 2 (2019-12-12) "same and energy sight" Space (shi 12) 2719 The # Toronist on the Statistical Computing		LOTA			144
fatfarr: xi6 64-m04-mirgu22/x64 (64-bit)		0.5704	454 rds of 55 cm	tables	
To Tree solitivane and comp with ACOULTELY NO SARDAITS.		Orazata tuesta	sta che, of so ye	lables .	
on are valcors to reciptribute 't under partain conditions.		values Lonelin	wals, of trearist	lics	
ge "Diewegt of Threadogt far distribution letails.		2 CH	we can, of sevent	241-24	
to a collaborative protect with now contributors.	III soto	100 cés: or 55 variables			
yes a serie Marten-11" for more tribured for and		0.6383	We add. 107 64- + 811		
ClearionG' on how so size a or a peckages in publications.		(1 est);	100 180, 07 55 18		
ype (secol) for some dense, "help()] for on-line help, or help, start() for an ere, browner interface to help.		() \$34 K	writes, of up said		
hilp.start() for an ene, browser interface to help. Mar (aC) in gain a		E 2344	99 sts. of 59 vari		
		U sat h	We disc of 3D early		
borkspace loaded from		D sald	100 pbs. of 90 ym		
		R.dr.	mille it manif	4.1	
		Rim. Nets. Relation			-
			a Antre Green		
		T & tray			
		A fastly		440.5	Manual Annahara
		AN222 N [2-10	01/29 2022 5 39 9 9
		C States		22.8	1+,37.3612 1.35.90
				7.2 68	6414-2020 D4K/04
		1 . R		44.05	10121-202-1527W
		D 8.		31,2112	141/20 2020 158 494
		in 10			
		1.1.1			
		22			
		() ()			
		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1			

*This is the R studio interface. The "Console" tab is where you type your commands and get R's feedback. But first, to make your future work easier, set up a work directory where you intend to store all related documents.

0 • 0 🕲 🚭 • 🕞 🚱 🖉 🖉	New Session		
Console Terminal × Jobs ×	Interrupt R Terminate R	5	Em
R version 3.6.2 (2019-1 Copyright (C) 2019 The	Restart R Ctrl+Shift+F10	uting	Dat
Platform: x86_64-w64-mi	Set Working Directory	To Source File Location	DF
R is free software and You are welcome to redi	Load Workspace	To Files Pane Location	D r
Type 'license()' or 'li	Save Workspace As	Choose Directory Ctrl+Shift+H	b
R is a collaborative pr	Clear Workspace		0 =
Type 'contributors()' f 'citation()' on how to	Quit Session Ctrl+Q ons .		
Type 'demo()' for some de	emos, 'help()' for on-line help	n or	0 9
'help.start()' for an HTM	L browser interface to help.		0 9
Type 'q()' to quit R.			0 9
[workspace loaded from ~/	(.RData]		0 9
>			0.
			File

Now the basic set-ups are done and you can start dealing with Twitter. Click in the console and start typing your commands. To finish and run a command, hit the Enter key.

Install package "rtweet" and load it

install.packages("rtweet")

library(rtweet)

Store all the API keys

api_key <- "..." api_secret_key <- "..." access_token <- "..." access_token_secret <- "..."

#replace the ... with corresponding keys and tokens.

Authenticate the connection

token <- create_token (
app = "NAMEOFYOURAPP", #use the name you entered for your app
consumer_key = api_key,
consumer_secret = api_secret_key,
access_token = access_token,
access_secret = access_token_secret)</pre>

This will create an environment in the workspace called "token", and you should see the values of your API keys in the "Environment" tab in the upper-right part of R studio. That means a connection between R and Twitter has been built successfully!

Environment History (Connections		
💣 🔒 📑 Import Dataset	• 1		List • 🛛 🕝 •
🐴 Global Environment 🔹		Q,	
🜔 sb7	99 obs. of 90 variables		-
🜔 sb8	100 obs. of 90 variables		
🜔 sb9	99 obs. of 90 variables		
token	Environment		Q,
D total	199 obs. of 90 variables		
🜔 total0	4757 obs. of 90 variables		
tweets	List of 200		Q,
tweets.df	200 obs. of 16 variables		
Values			
access_token			
access_token_sec	0		
api_key			
api_secret_key			

Extract relevant tweets (regular approaches)

To search for tweets that are relevant to a topic (e.g. according to hashtags)

DF <-*search_tweets* (*q="YOURKEYWORD"*, *n=NUMBER OF TWEETS*)

You'll have a data frame that contains the requested data. But only recent tweets (6-9 days) can be extracted with the standard API product! To extract all it is NECESSARY to update to a more advanced API package. See *advanced approaches*. You can click on the data frame in the "Environment tab" to view the dataset.

Alternatively, to extract tweets and replies from a certain account

DF<-get_timeline(q="@SCREENNAME", n=NUMBER OF TWEETS)

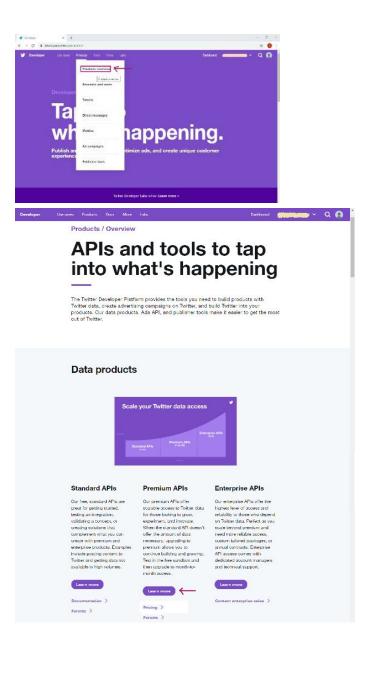
Note that you need to search by the screen name (the one starting with @), NOT the page name.

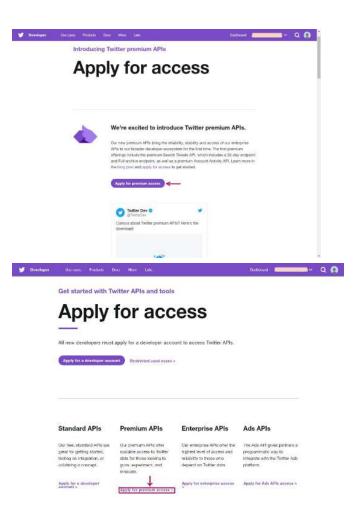
You can set the n to a larger number to make it convenient. However, there's an upper limit of 3,200 tweets.

If you want to extract historical comments rather than tweets, there's no direct way to do that with regular approaches. But you can turn to *advanced approaches* in the next section.

Extract comments (advanced approaches)

The regular approaches are handy, but there are obvious limitations when you want to extract comments. The first approach does give you a sheet containing comments, but it has a strict time restriction; the second approach does not allow you to see comments from other users besides the account of interest. Therefore, you may want to upgrade to an advanced developer toolkit. **API Sandbox** can be a good choice because it's FREE (bad news is that it requires more manual input if you expect to have large data). You do this by going to your developer's webpage and click on the *Products* tab. The Sandbox has two versions: 30-Days and full archive. Choose one (or both) according to your needs and give the new dev environment(s) a label (random name).





Now that you have chosen your product(s), you can now see what they will enable you to do on your developer's dashboard (<u>https://developer.twitter.com/en/dashboard</u>). This is also where you'll see the usage of your products.

To activate a product, set up a dev environment and label it.

🗅 🛓 deskipe bollener verskaldener			-	4
Assistant Gaussian Folloch Data Mon	3469		Distant -	- (
ee.est				
March receiption of a particular departed with a particular product of the particular department of the particular department department of the particular department of the particular de	(alones) (daplicate) (dap posed) (ext	the analysis and solution	the latter of the product of	6.850 (H) #1 (H) #4 (J
(Bevit	Search the month	Projected respirate 0	Toppy the monity	Projected to rate
rch Tweets: 30 Days / Sandbax ***	Requests this recent 0 of 250 TWC/IS states	0	D of 25k	Projected livests 0
Bruse:	TWEETS ESHE			
nutree BN	1	1-		
	\ <u>`</u>	/	100.003	
in na le <mark>rtin de la c</mark> elar i accio a adoi	ar ant also dag scale.			
rch Tweets: Full Archive / Sandbax ***	Departs this events 30 of 50	Frequencies 52	facets this result 4.50k of 5k	fodielaitacea S.OBk
ISTS (KAST	TWEETS ESHOR			
			1	-
TES	100%			- 1A
(aray			ORNO	
Devloper Decision Products Deco More	Later		Dathboard 🔚	<u> </u>
nebumente				
in using one of our substration APIs, you read to set up a dealers dent of the default of your packages, which can be used to police o	were the the endpoint. All of the sub	scipheri API packaget su	iportatient filles envito	viveres (number I).
dent on the details of your packages, which can be used to sporte a at disservance's appl.	nender verser som nakere de omstelleren mede	n pere systeme, inscitt if a	NE 1998 (10, 1925, 2005), 5	a magaine your
earch Tweets: 30-Days / Sandbox				1
This devery intromment has not been art up.			Continue	the second second
			Carlouin .	A STATE OF A
		Approxim		*****
earch Tweets Full Archive / Sandbox Deconverses later Jee Rane	-	Anti-seaso	. @	
earch Tweets Full Archive / Sandbox Deconverses later Jee Rane		Approxim	. C	
earch Tweets: Full Archive / Sandbox		742 Harrison	(24 (24)	de reaction and a
earch Tweets: Full Archive / Sandbox In revieween later account Activity API / Sandbox		App search	E CA	
earch Tweets: Full Archive / Sandbox In reminimum Are Bare		Agrissio		der resservensend
earch Tweets: Full Archive / Sandbox In reminimum Are Bare	-	App source		No many consents
earch Tweets: Full Archive / Sandbox In reminimum Are Bare	-	Page senses		N minimum
earch Tweet's Full Archive / Sandbox The performance test coount Archivity API / Sandbox where examples are the null been set up.			(4) (4)	No reactions
erch Tweets Full Archive / Sandbox Inversement Ideal Are Run-				Ale veze contractor
earch Tweet's Full Archive / Sandbox The performance test coount Archivity API / Sandbox where examples are the null been set up.				
each Tweet's Full Archive : Sandbox The oreinweet later coount Archivey API / Sandbox Set up Search Tweets: 30- Dev environment label			. (*	NY TOLE HANDING
arch Tweeter Ful Archive ; Sandbare In wetweet Later Argelium count Activity API / Sandbare er er of The for example of the relation of the set of the Set up Search Tweets: 30-				A real and a
arch Tweeter Ful Archive : Sandbox Investment Ltdr Arg Harr count Activity API / Sandbox of pro- The for examplement har rid from set up. Set up Search Tweets: 30- Dev environment label				
arch Tweeter Ful Archer / Sandbare In universited taid count Arthily API / Sandbare of other The for example of the end tweet set up Set up Search Tweets: 30- Dev environment label BeCreativeHere 1				
arch Tweeter Ful Archive : Sandbox Investment and Archive : Sandbox count Archivey API / Sandbox Set up Search Tweets: 30- Dev environment label BeCreativeHere 1 App				
arch Tweeter Ful Archive : Sandbox Investment and Archive : Sandbox count Archivey API / Sandbox Set up Search Tweets: 30- Dev environment label BeCreativeHere 1 App				

The above example is based on Sandbox 30-Days, but you can choose according to your needs and budget. Sandbox is a premium offer, albeit free.

After completing the setup, you will be redirected to the Dashboard and you will notice that the line in red "you must first set up a dev environment..." under your selected product is gone. This means that you have successfully activated the product and are ready to proceed to advanced extraction!

Search through the entire history (full archive)

DF<- search_fullarchive (q="@SCREENNAME", n=100, fromDate = YYYYMMDDHHMM, toDate=YYYYMMDDHHMM, env_name = "YOURLABEL")

you can request data 50 times per month, 100 tweets per request, meaning that n has to be no larger than 100.

you don't always need to havethe fromDate term in your command, especially when you're going backwards into history from a later date for 100 tweets.

Search through the past month (30-Days)

DF<- search_30day (q="@SCREENNAME", n=100, fromDate = YYYYMMDDHHMM, toDate=YYYYMMDDHHMM, env_name = "YOURLABEL")

You can request data 250 times per month, 100 tweets per request.

Again, n<=100 and fromDate isn't always necessary.

This only gives you data from the past month, and the full archive option can do the same for you. But it's useful when you need past month's data urgently and you know there were more than 5000 tweets where the page was tagged or replied to.

Note that the data you get will contain 1) comments under the page's tweets, 2) retweets of the page's tweets and 3) tweets created by other users where the page is tagged. However, it's not hard to tell which ones are comments and filter out the rest (see *Filtering* section for details).

With the upper limits for the number of tweets you may get per request though, you might have to extract multiple times when there are too many tweets. In that case, you can merge all the datasets into a big, neat one. Here is one of the simple ways to do that.

FINALDATA<-rbind(DF1, DF2,..., DFn)

Transfer data to a csv file

install package "data.table" and load it

Install.packages("data.table")

library(data.table)

use the "fwrite" function

fwrite(FINALDATA, file = "CSVNAME.csv")

This is one way to do things. Regular way of write.csv does not work because the data frame contains list variables.

Alternative way (less efficient):

sapply(DF,class) #see which variables are list-type

unlist(DF\$COLUMNNAME) #unlist those variables **OR**

toString(DF\$COLUMNNAME) #transfer the variables to string-type

write.csv(DF,"CSVNAME.csv")

The csv file will be automatically saved to your work directory. Now you have your data!