

# Text Mining using Twitter Data

# Contents

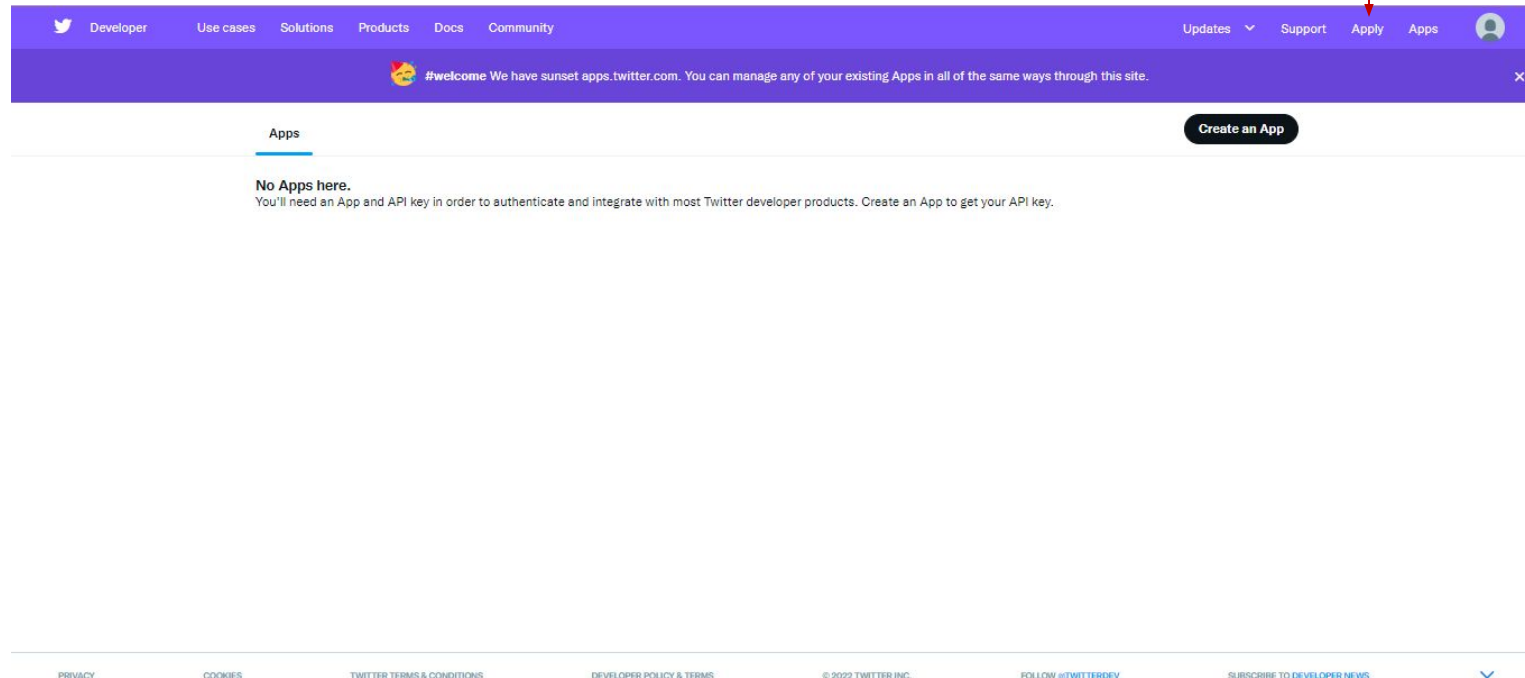
1. Why is Twitter Data Useful
2. Creating Twitter developer account
3. Creating App
4. Text mining in Python
  - Authentication of Twitter Account in Python
  - Fetching data from Twitter using 'tweepy' library
  - Cleaning data for text analysis
  - Generating WordCloud
  - Sentiment Analysis

# Why is Twitter Data Useful

- Text mining is getting a lot attention in last few years, due to an exponential increase in digital text data from web pages and social media services.
- Twitter data constitutes a rich source that can be used for capturing information about any topic. This data can be used for finding trends related to a specific keyword, measuring brand sentiment, and gathering feedback about new products and services etc.

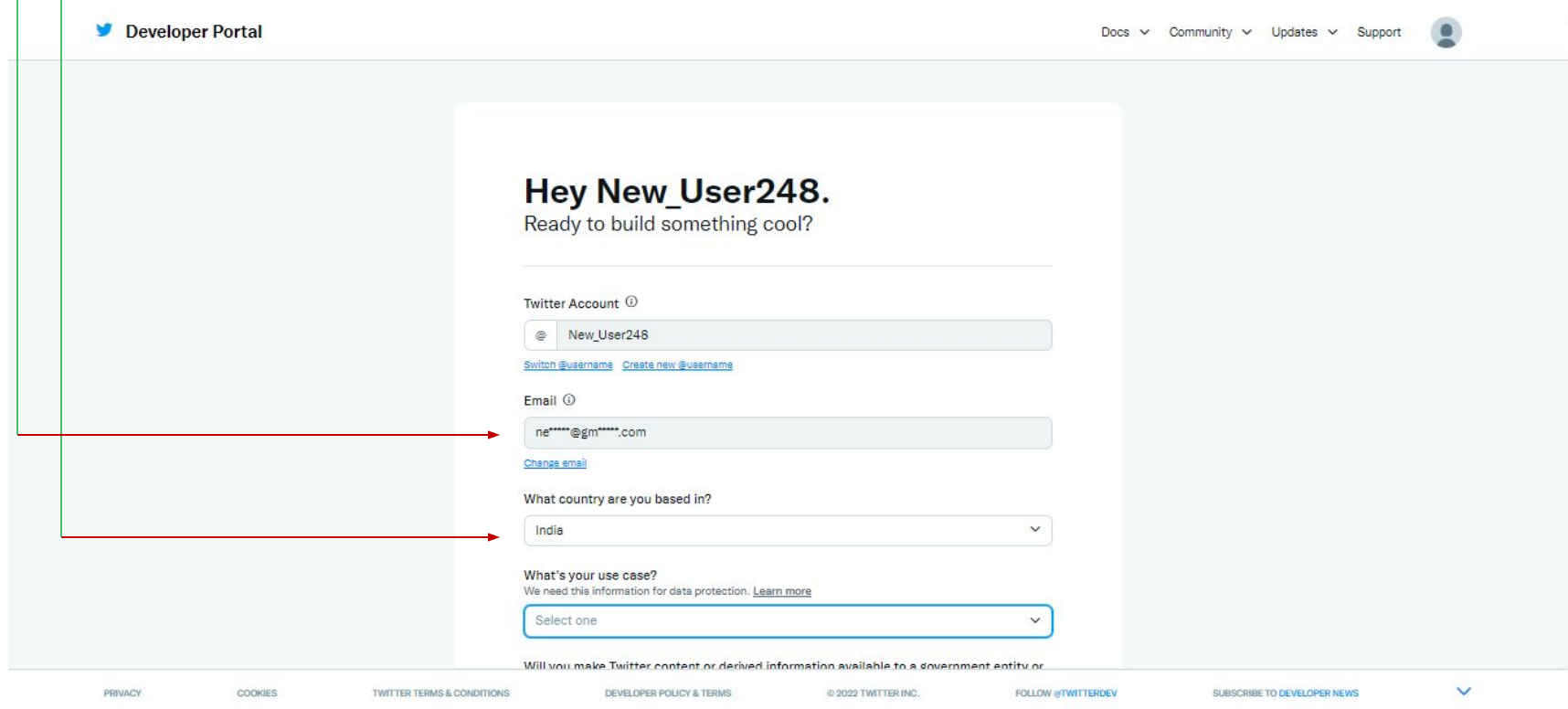
# Creating Twitter Developer Account

- Twitter account is required to fetch the data from it. If not, first create a Twitter account.
- Keep the Twitter account running on one webpage.
- Open another webpage of [apps.twitter.com](https://apps.twitter.com)
- Click on the **apply**.



# Creating Twitter Developer Account

- Fill in your verified email.
- Select the **Country**.
- If you already have developer account, you can go to slide number 24.



The screenshot shows the Twitter Developer Portal account creation interface. At the top, the header includes the Twitter logo, 'Developer Portal', and navigation links for 'Docs', 'Community', 'Updates', and 'Support'. The main content area is titled 'Hey New\_User248. Ready to build something cool?'. Below this, there are four input fields: 'Twitter Account' (with a dropdown showing '@ New\_User248'), 'Email' (with a dropdown showing 'ne\*\*\*\*\*@gm\*\*\*\*\*.com'), 'What country are you based in?' (with a dropdown showing 'India'), and 'What's your use case?' (with a dropdown showing 'Select one'). A green line with arrows points from the list items to the 'Email' and 'Country' fields. At the bottom, there is a footer with links for 'PRIVACY', 'COOKIES', 'TWITTER TERMS & CONDITIONS', 'DEVELOPER POLICY & TERMS', '© 2022 TWITTER INC.', 'FOLLOW @TWITTERDEV', and 'SUBSCRIBE TO DEVELOPER NEWS'.

Developer Portal Docs Community Updates Support

Hey New\_User248.  
Ready to build something cool?

Twitter Account ⓘ  
@ New\_User248  
[Switch @username](#) [Create new @username](#)

Email ⓘ  
ne\*\*\*\*\*@gm\*\*\*\*\*.com  
[Change email](#)

What country are you based in?  
India

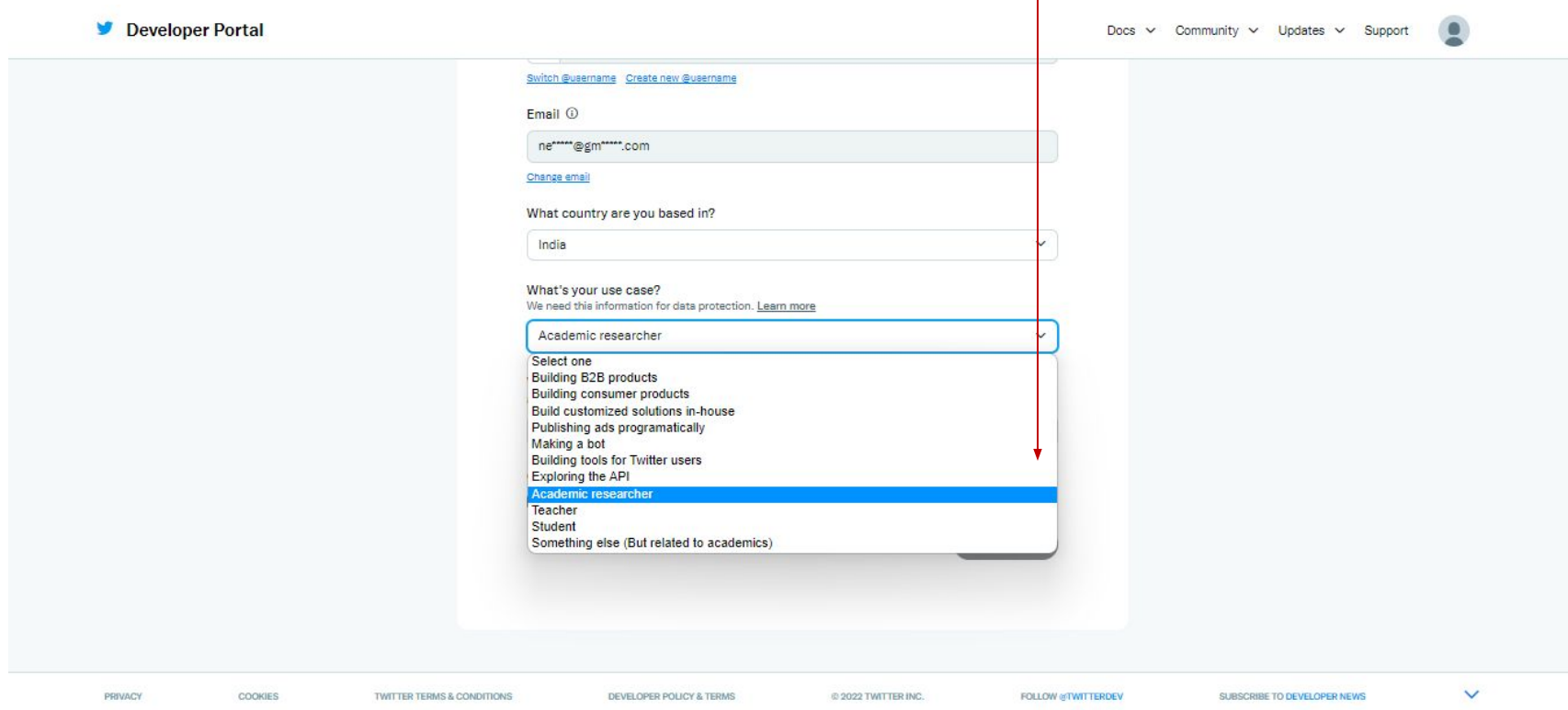
What's your use case?  
We need this information for data protection. [Learn more](#)  
Select one

Will you make Twitter content or derived information available to a government entity or

PRIVACY COOKIES TWITTER TERMS & CONDITIONS DEVELOPER POLICY & TERMS © 2022 TWITTER INC. FOLLOW @TWITTERDEV SUBSCRIBE TO DEVELOPER NEWS

# Creating Twitter Developer Account

- Click on Academic Researcher.



The screenshot shows the Twitter Developer Portal account creation interface. At the top, the 'Developer Portal' logo is on the left, and navigation links for 'Docs', 'Community', 'Updates', and 'Support' are on the right. The main form area contains the following fields:

- [Switch @username](#) and [Create new @username](#) links.
- Email** field with a help icon, containing the text 'ne\*\*\*\*\*@gm\*\*\*\*\*.com'. Below it is a [Change email](#) link.
- What country are you based in?** dropdown menu with 'India' selected.
- What's your use case?** dropdown menu. The text below it reads: 'We need this information for data protection. [Learn more](#)'. The dropdown is open, showing a list of options: 'Academic researcher' (selected), 'Select one', 'Building B2B products', 'Building consumer products', 'Build customized solutions in-house', 'Publishing ads programmatically', 'Making a bot', 'Building tools for Twitter users', 'Exploring the API', 'Academic researcher' (highlighted in blue), 'Teacher', 'Student', and 'Something else (But related to academics)'.

The footer contains links for 'PRIVACY', 'COOKIES', 'TWITTER TERMS & CONDITIONS', 'DEVELOPER POLICY & TERMS', '© 2022 TWITTER INC.', 'FOLLOW @TWITTERDEV', and 'SUBSCRIBE TO DEVELOPER NEWS'.

# Creating Twitter Developer Account

- Review and accept the developer agreement.
- Scroll down and tick the box. Click on **Submit**.

Developer Portal

Docs Community Updates Support

## Developer agreement & policy

### Developer Agreement

Effective: March 10, 2020

This Twitter Developer Agreement (“**Agreement**”) is made between you (either an individual or an entity, referred to herein as “**you**”) and Twitter (as defined below) and governs your access to and use of the Licensed Material (as defined below). Your use of Twitter’s websites, SMS, APIs, email notifications, applications, buttons, embeds, ads, and our other covered services is governed by our general Terms of Service and Privacy Policy.

PLEASE READ THE TERMS AND CONDITIONS OF THIS AGREEMENT CAREFULLY, INCLUDING ANY LINKED TERMS REFERENCED BELOW, WHICH ARE PART OF THIS LICENSE AGREEMENT. BY USING THE LICENSED MATERIAL, YOU ARE AGREEING THAT YOU

☒ Accept Terms & Conditions

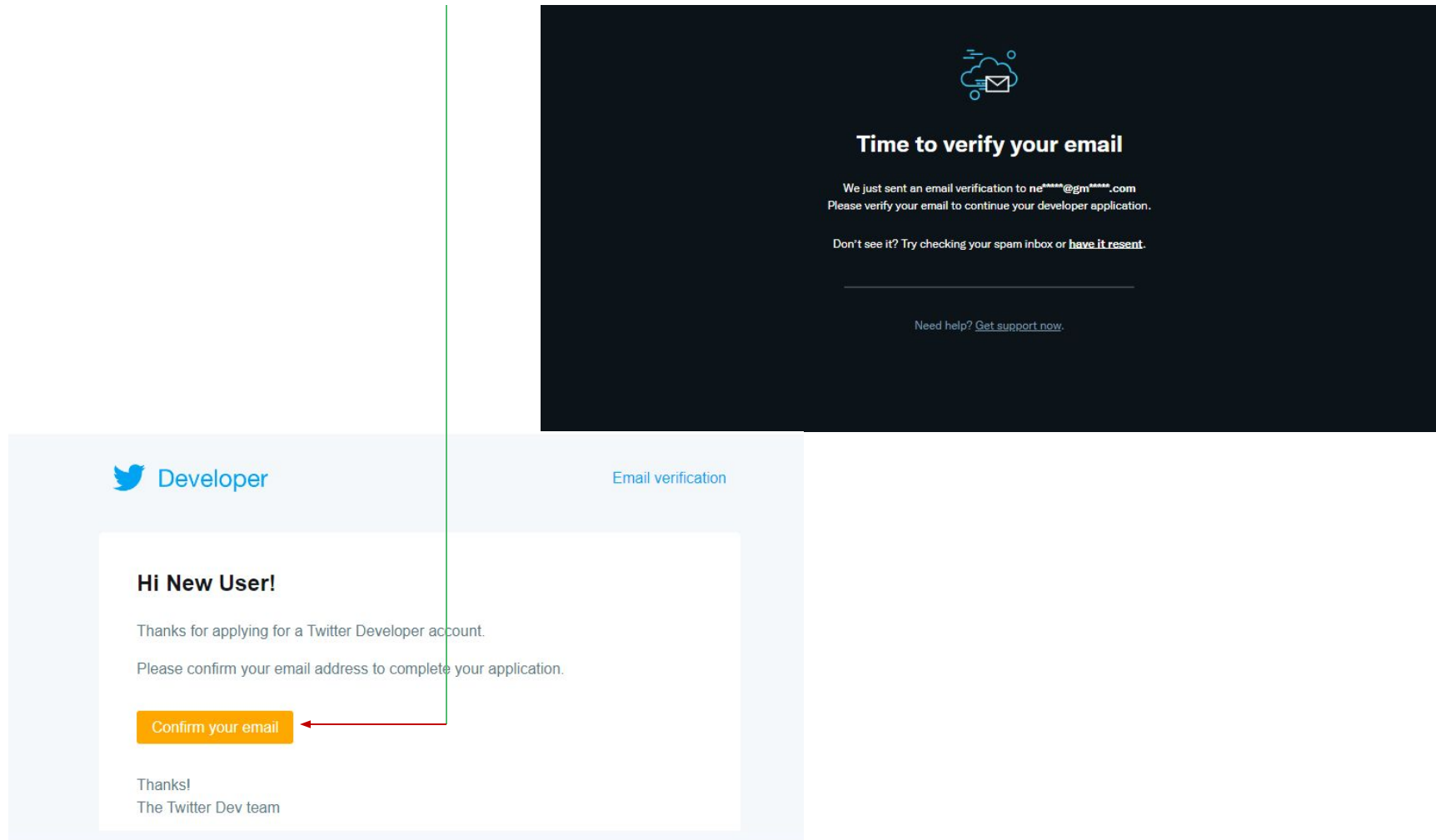
By clicking on the box, you indicate that you have read and agree to this [Developer Agreement](#) and the [Twitter Developer Policy](#), additionally as it relates to your display of any of the Content, the [Display Requirements](#), as it relates to your use and display of the Twitter Marks, the [Twitter Brand Assets and Guidelines](#), and as it relates to taking automated actions on your account, the [Automation Rules](#). These documents are available in hardcopy upon request to Twitter.

Back Submit

PRIVACY COOKIES TWITTER TERMS & CONDITIONS DEVELOPER POLICY & TERMS © 2022 TWITTER INC. FOLLOW @TWITTERDEV SUBSCRIBE TO DEVELOPER NEWS

# Creating Twitter Developer Account

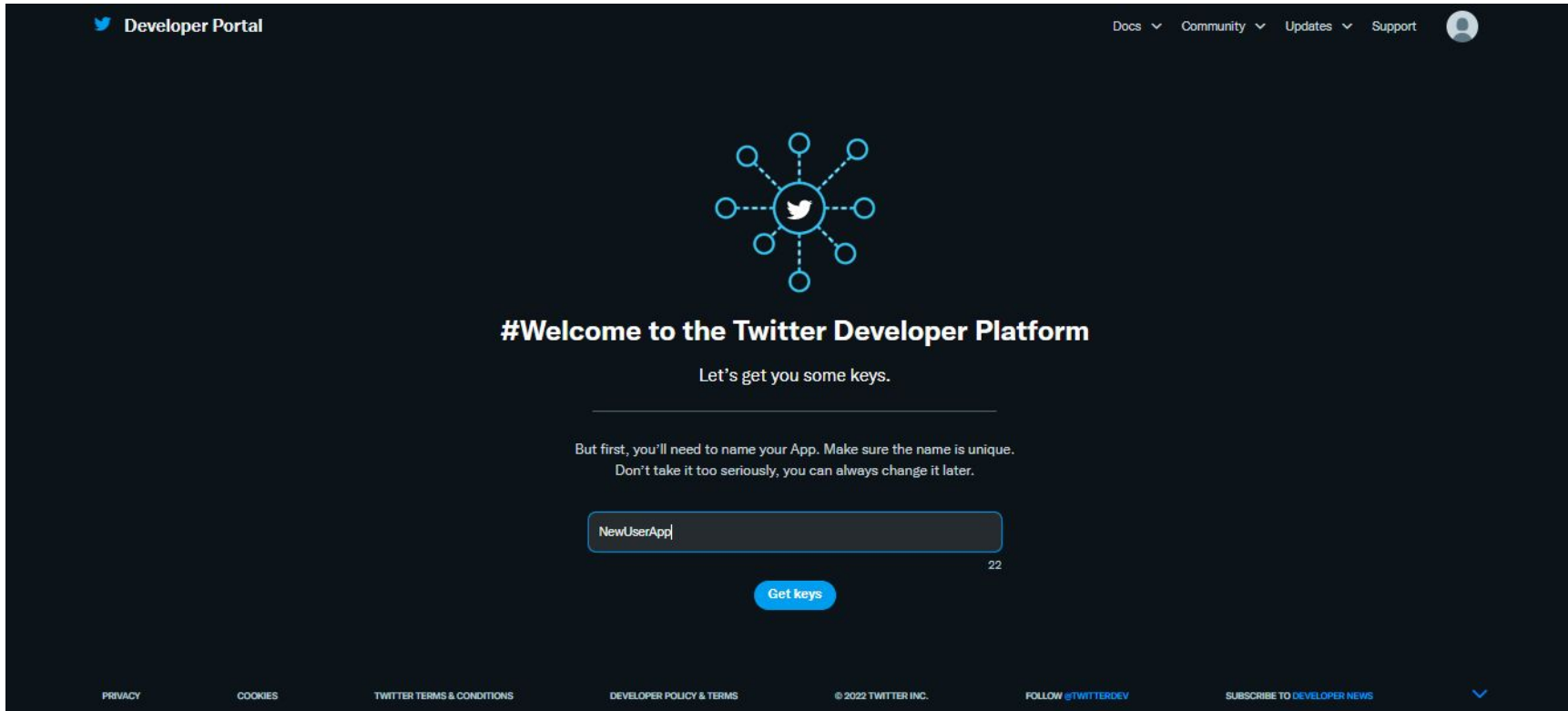
- Once this page is opened, open the email received to registered email id and click on the confirm your email.





# Creating App

- Once you click on the **confirm your email**, it will direct to this page.
- Give an **App Name**. (It should be UNIQUE for each developer).
- Once done, click on **Get Keys**.



The screenshot shows the Twitter Developer Portal interface for creating a new app. At the top, the header includes the Twitter logo, 'Developer Portal', and navigation links for Docs, Community, Updates, and Support, along with a user profile icon. The main content area features a central graphic of a Twitter bird surrounded by a network of nodes. Below this, the heading '#Welcome to the Twitter Developer Platform' is displayed, followed by the instruction 'Let's get you some keys.' and a horizontal line. A message states: 'But first, you'll need to name your App. Make sure the name is unique. Don't take it too seriously, you can always change it later.' A text input field contains the text 'NewUserApp'. To the right of the input field is a small '22' character count. Below the input field is a blue 'Get keys' button. The footer contains links for Privacy, Cookies, Twitter Terms & Conditions, Developer Policy & Terms, Copyright 2022 Twitter Inc., Follow @TWITTERDEV, and Subscribe to Developer News, followed by a downward arrow icon.

Developer Portal

Docs Community Updates Support

#Welcome to the Twitter Developer Platform

Let's get you some keys.

But first, you'll need to name your App. Make sure the name is unique.  
Don't take it too seriously, you can always change it later.

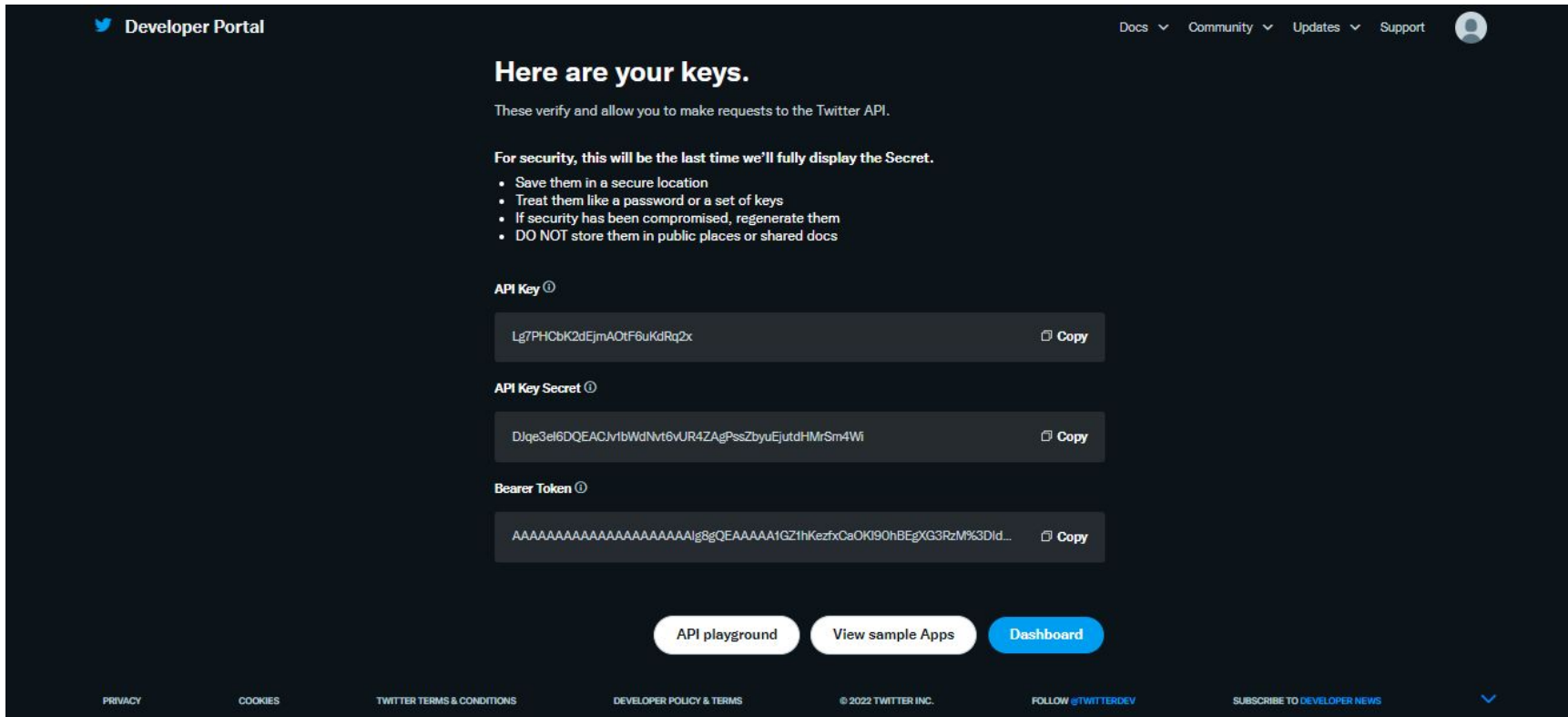
NewUserApp 22

Get keys

PRIVACY COOKIES TWITTER TERMS & CONDITIONS DEVELOPER POLICY & TERMS © 2022 TWITTER INC. FOLLOW @TWITTERDEV SUBSCRIBE TO DEVELOPER NEWS

# Creating App

- Save the **API Key** and **API Secret Key** into a notepad/external file as they are required for further analysis and will not be available later.
- Once copied, click on **Dashboard**.



The screenshot shows the Twitter Developer Portal interface. At the top, there's a navigation bar with the Twitter logo, 'Developer Portal', and links for Docs, Community, Updates, and Support. The main heading is 'Here are your keys.' followed by a subtext: 'These verify and allow you to make requests to the Twitter API.' Below this, a security warning states: 'For security, this will be the last time we'll fully display the Secret.' and lists four instructions: 'Save them in a secure location', 'Treat them like a password or a set of keys', 'If security has been compromised, regenerate them', and 'DO NOT store them in public places or shared docs'. The interface then displays three key types with their respective values and 'Copy' buttons: 'API Key' (Lg7PHCbK2dEjmAOtF6uKdRq2x), 'API Key Secret' (DJqe3el6DQEAQJvrbWdNvt6vUR4ZAqPssZbyuEjtdHMrSm4Wi), and 'Bearer Token' (AAAAAAAAAAAAAAAAAAAAAg8gQEAAAAA1GZthKefzxCsOKI90hBEgXG3RzM%3Dld...). At the bottom, there are three buttons: 'API playground', 'View sample Apps', and 'Dashboard'. The footer contains links for Privacy, Cookies, Twitter Terms & Conditions, Developer Policy & Terms, Copyright 2022 Twitter Inc., Follow @TWITTERDEV, and Subscribe to Developer News.

Developer Portal

Docs Community Updates Support

## Here are your keys.

These verify and allow you to make requests to the Twitter API.

**For security, this will be the last time we'll fully display the Secret.**

- Save them in a secure location
- Treat them like a password or a set of keys
- If security has been compromised, regenerate them
- DO NOT store them in public places or shared docs

**API Key**

Lg7PHCbK2dEjmAOtF6uKdRq2x [Copy](#)

**API Key Secret**

DJqe3el6DQEAQJvrbWdNvt6vUR4ZAqPssZbyuEjtdHMrSm4Wi [Copy](#)

**Bearer Token**

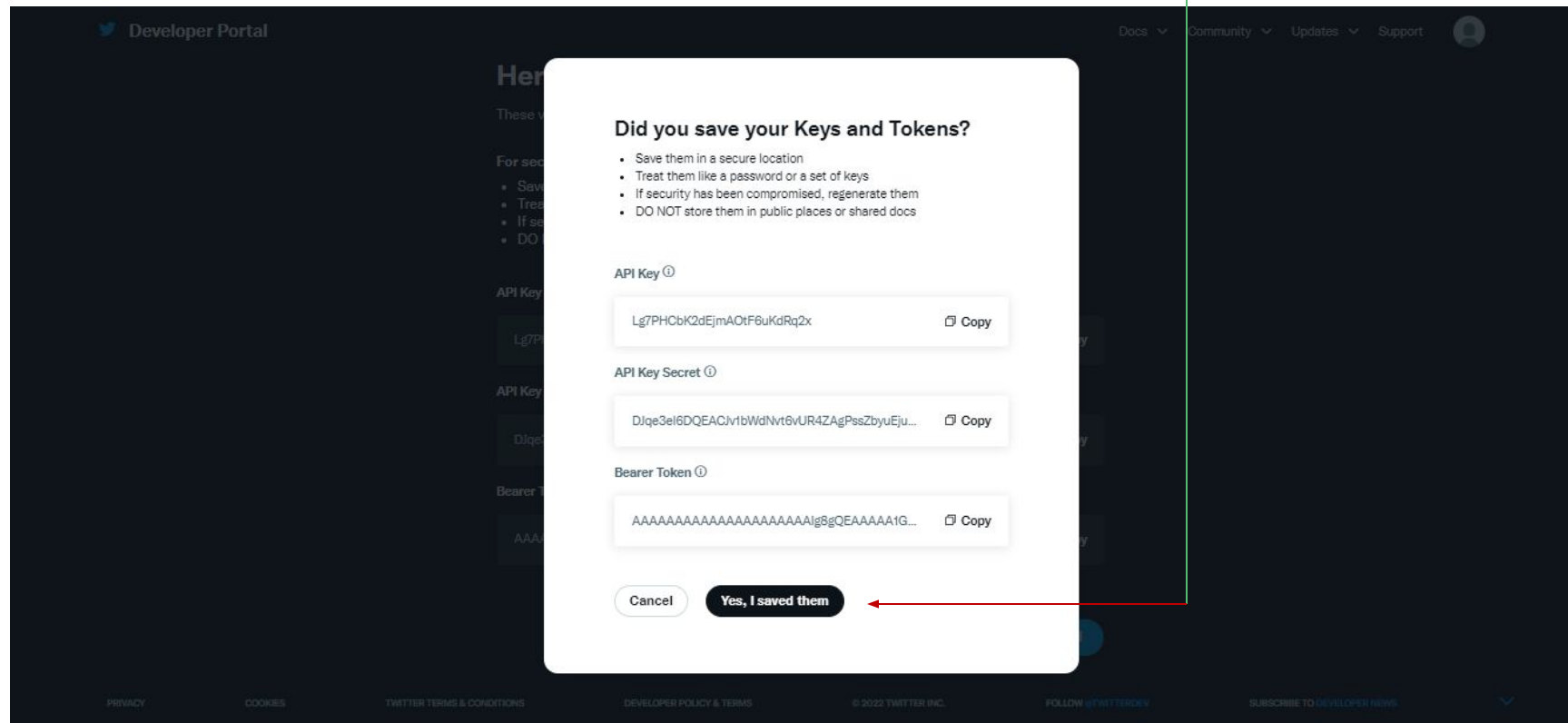
AAAAAAAAAAAAAAAAAAAAAg8gQEAAAAA1GZthKefzxCsOKI90hBEgXG3RzM%3Dld... [Copy](#)

[API playground](#) [View sample Apps](#) [Dashboard](#)

PRIVACY COOKIES TWITTER TERMS & CONDITIONS DEVELOPER POLICY & TERMS © 2022 TWITTER INC. FOLLOW @TWITTERDEV SUBSCRIBE TO DEVELOPER NEWS

# Creating App

- It will load a popup form to confirm. Click on Yes, I saved them.



# Elevated Access

- Now that the app is created, essential access is granted. But in order to extract tweets, elevated access is required.

- From the side menu, go to **Products > Twitter API v2**.

- Switch to **Elevated** tab and click **Apply**.


The screenshot displays the Twitter Developer Portal interface. On the left, a dark sidebar contains the 'Developer Portal' logo and a menu with 'Dashboard', 'Projects & Apps', 'Products' (marked with a 'NEW' badge), and 'Twitter API v2' (which is highlighted with a red arrow). The main content area is titled 'Twitter API v2' and features three tabs: 'Essential', 'Elevated' (selected), and 'Academic Research'. Below the tabs, the 'Elevated' section is shown, including an 'Overview' with the text 'Higher levels of access to the Twitter API for free with an approved application.' and a table of features:

Apps	3 environments per project
Tweets	2M Tweets per month / Project
Cost	free

Below the table, a question 'Do you need Elevated access for your Project?' is followed by an 'Apply' button, which is pointed to by a red arrow. At the bottom, there is a section for 'Elevated features' with a sub-section for 'Tweets'. The footer contains links for 'PRIVACY', 'COOKIES', 'TWITTER TERMS & CONDITIONS', 'DEVELOPER POLICY & TERMS', '© 2022 TWITTER INC.', 'FOLLOW @TWITTERDEV', and 'SUBSCRIBE TO DEVELOPER NEWS'.

# Elevated Access

- Double check the **username** and **email id** and if correct, proceed to the rest of the form.

 **Developer Portal**

### #TheBasics

To start, you'll need to confirm your Twitter @username and the email associated with it. You'll use them to log in to the Developer Platform.

### Team accounts

For accounts that need multiple users, create a [team developer account](#) instead.


1 Basic info

2 Intended use

3 Review

4 Terms

All fields are required unless marked optional



New\_User248  
@New\_User248

[Switch @username](#)

This @username will be used to log in to your account.

ne\*\*\*\*@gm\*\*\*\*.com

[Change email address](#)

This will be used for communications about the application status, and will be used throughout the entire developer access process. [Learn more](#)

Individual developer account

[Switch to a team developer account](#)

You are signing up for an individual developer account. [Learn the differences between team & individual developer account.](#)

Back

Next

PRIVACY

COOKIES


TWITTER TERMS & CONDITIONS

DEVELOPER POLICY & TERMS

© 2022 TWITTER INC.

[FOLLOW @TWITTERDEV](#)

[SUBSCRIBE TO DEVELOPER NEWS](#)



# Elevated Access

- Double check your name and country.
- Select **Some experience** for the question what's your current coding skill level?
- Once all fields are verified, click **Next**.

The screenshot shows the Twitter Developer Portal sign-up process. On the left is a dark sidebar with the Twitter logo and 'Developer Portal' text. Below it, under '#TheBasics', is a paragraph: 'To start, you'll need to confirm your Twitter @username and the email associated with it. You'll use them to log in to the Developer Platform.' Further down, under 'Team accounts', is another paragraph: 'For accounts that need multiple users, create a [team developer account](#) instead.'

The main content area has a top navigation bar with links: Docs, Community, Updates, and Support, followed by a user profile icon. Below this is a progress indicator with four steps: 1 Basic info (active), 2 Intended use, 3 Review, and 4 Terms.

The 'Basic info' section contains three main fields:

- Individual developer account:** A heading with a link 'Switch to a team developer account'. Below it is a text input field with 'New\_User248' and a note: 'This is permanent and can't be changed.'
- What country are you based in?:** A dropdown menu showing 'India'.
- What's your current coding skill level?:** A dropdown menu with 'Some experience' selected. A red arrow points to this dropdown from the text 'Select Some experience for the question what's your current coding skill level?' in the list above.

At the bottom of the form is a checkbox labeled 'Want updates? (optional)' with the text 'Don't miss the latest news and tips emailed to you.' and the option 'Yes, send updates.'.

At the bottom right of the form are two buttons: 'Back' and 'Next'. A red arrow points from the 'Next' button to the text 'Once all fields are verified, click Next.' in the list above.

The footer contains links: PRIVACY, COOKIES, TWITTER TERMS & CONDITIONS, DEVELOPER POLICY & TERMS, © 2022 TWITTER INC., FOLLOW @TWITTERDEV, and SUBSCRIBE TO DEVELOPER NEWS, followed by a chevron icon.

# Elevated Access

- Write the description for each question under How will you use the Twitter API or Twitter data? Minimum number of characters for answers are specified for each question.

- You can take help from this link:

<https://www.jcchouinard.com/apply-for-a-twitter-developer-account/>

The screenshot shows the Twitter Developer Portal interface. On the left is a dark sidebar with the 'Developer Portal' logo and several sections: '#ImportantStuff' (explaining the application's purpose and Twitter's goal), 'What's not allowed' (listing prohibited activities like surveillance), 'Automation guidelines' (warning about automated activity), and 'Provide thorough answers' (emphasizing the need to understand the platform). The main content area is titled 'Intended use' and contains the question 'How will you use the Twitter API or Twitter Data?'. Below this, it says 'In your words' and provides instructions to describe the intended use of Twitter data and/or APIs. A large text input box is provided for the answer, with a placeholder text 'Please be thoughtful and thorough' and a character count of 200. At the bottom of the form are 'Back' and 'Next' buttons. The footer contains links for Privacy, Cookies, Twitter Terms & Conditions, Developer Policy & Terms, Copyright 2022 Twitter Inc., Follow @TWITTERDEV, and Subscribe to Developer News.

Developer Portal

#ImportantStuff

This part of the application helps us make sure that our data will be handled in accordance with Twitter's Developer Policies.

Our ultimate goal is to keep Twitter a safe and healthy space for public conversation.

What's not allowed

Some activities (like surveillance) are never allowed on Twitter. Take a look at our restricted uses page. Double-check that your intended use is policy compliant.

Automation guidelines

If you plan on enabling any sort of automated activity on the platform (like a bot), be sure to review our automation rules.

Provide thorough answers

We need to fully understand your

Docs Community Updates Support

1 Basic info 2 Intended use 3 Review 4 Terms

All fields are required unless marked optional

How will you use the Twitter API or Twitter Data?

In your words

In English, please describe how you plan to use Twitter data and/or APIs. The more detailed the response, the easier it is to review and approve. [Learn what specific information to include in your use case](#)

Please be thoughtful and thorough

200

Back Next

PRIVACY COOKIES TWITTER TERMS & CONDITIONS DEVELOPER POLICY & TERMS © 2022 TWITTER INC. FOLLOW @TWITTERDEV SUBSCRIBE TO DEVELOPER NEWS

Docs

Communi

1 Basic info

2 Intended use

3 Review

4 Terms

### The specifics

Please answer each of the following with as much detail and accuracy as possible. Failure to do so could result in delays to your access to Twitter developer platform or rejected applications. Need help? [Get support now](#).

Are you planning to analyze Twitter data?

✓

Yes

Please describe how you will analyze Twitter data including any analysis of Tweets or Twitter users.

Please be thoughtful and thorough

100

Will your App use Tweet, Retweet, Like, Follow, or Direct Message functionality?

✓

Yes

Docs

Community

1 Basic info

2 Intended use

3 Review

4 Terms

Will your App use Tweet, Retweet, Like, Follow, or Direct Message functionality?

✓

Yes

Please describe your planned use of these features.

Please be thoughtful and thorough

100

Do you plan to display Tweets or aggregate data about Twitter content outside Twitter?

✓

Yes

Please describe how and where Tweets and/or data about Twitter content will be displayed outside of Twitter.

Please be thoughtful and thorough

TWITTER TERMS & CONDITIONS

DEVELOPER POLICY & TERMS

© 2022 TWITTER INC.

FOLLOW @TWITTERDEV

SUB

XOKIES

TWITTER TERMS & CONDITIONS

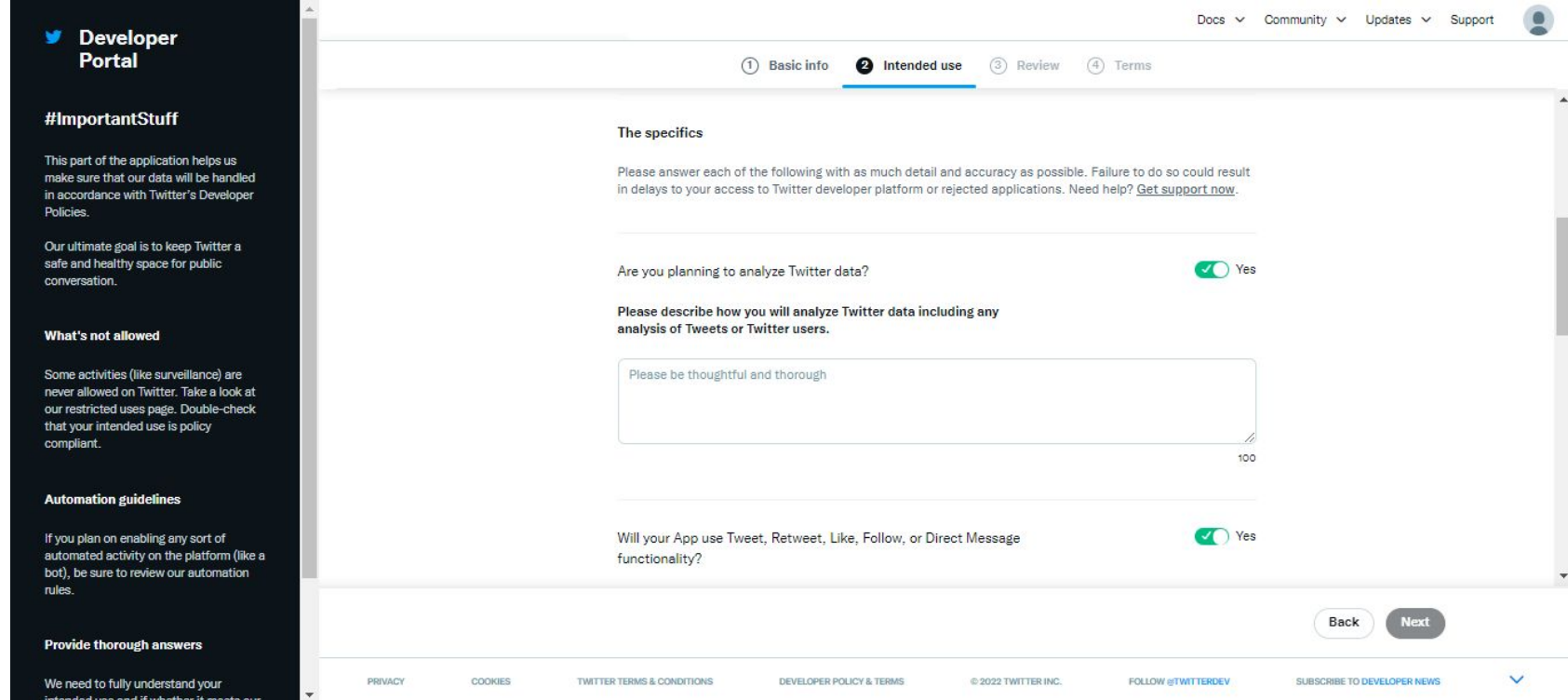
DEVELOPER POLICY & TERMS

© 2022 TWITTER INC.

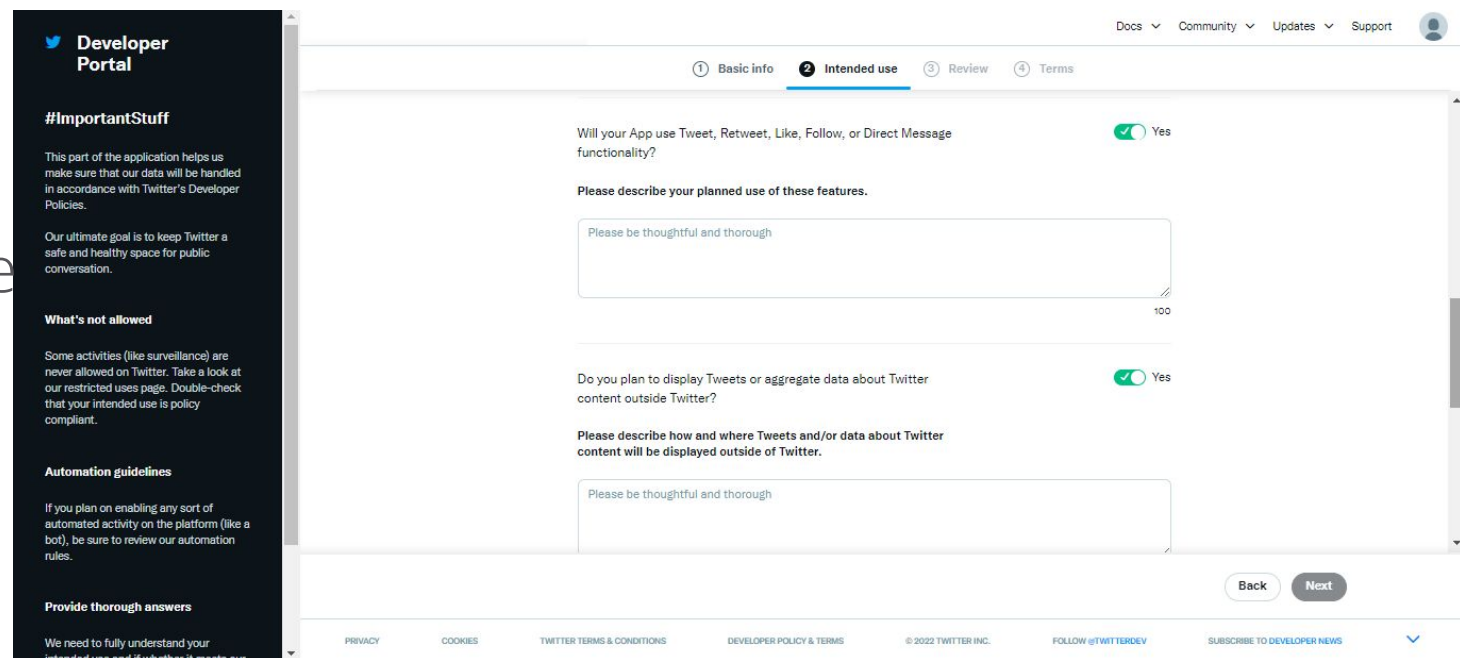
FOLLOW @TWITTERDEV

SUB





In case, the images in the above slide need a different layout.



# Elevated Access

- Double check the inputs and if correct, click **Next**. Else click **Edit**.

**Developer Portal**

**#Review**

Before moving forward, take a moment to look over your responses. If everything looks good, you can move on to the next step.

One thing to remember, the email address you provided will be used to contact you about your account.

1 Basic info 2 Intended use 3 Review 4 Terms

**Basic info** Edit

USE CASE Doing academic research	ACCOUNT NAME New_User248
ACCOUNT TYPE Individual developer account	TWITTER @HANDLE @New_User248
EMAIL ADDRESS ne****@gm****.com	WHAT COUNTRY DO YOU LIVE IN? India
CURRENT CODING SKILL Some experience	RECEIVE UPDATES ABOUT TWITTER API No

**Intended use** Edit

Back Next

PRIVACY COOKIES TWITTER TERMS & CONDITIONS DEVELOPER POLICY & TERMS © 2022 TWITTER INC. FOLLOW @TWITTERDEV SUBSCRIBE TO DEVELOPER NEWS

# Elevated Access

- Review and accept the Developer agreement & policy.
- Scroll down and tick the box. Click on Submit.

**Developer Portal**

**#Read&Accept**

We've carefully crafted our developer terms to make it readable and accessible. Our aim is to have a healthy and open platform for all.

Once you've read it and agreed, check the box below the agreement and then hit the submit application button on the bottom right.

Docs ▾ Community ▾ Updates ▾ Support ▾

① Basic info ② Intended use ③ Review ④ **Terms**

## Developer agreement & policy

### Developer Agreement

Effective: March 10, 2020

This Twitter Developer Agreement (“**Agreement**”) is made between you (either an individual or an entity, referred to herein as “**you**”) and Twitter (as defined below) and governs your access to and use of the Licensed Material (as defined below). Your use of Twitter's websites, SMS, APIs, email notifications, applications, buttons, embeds, ads, and our other covered services is governed by our general Terms of Service and Privacy Policy.

PLEASE READ THE TERMS AND CONDITIONS OF THIS AGREEMENT CAREFULLY, INCLUDING ANY LINKED TERMS REFERENCED BELOW, WHICH ARE PART OF THIS LICENSE AGREEMENT. BY USING THE LICENSED MATERIAL, YOU ARE AGREEING THAT YOU HAVE READ, AND THAT YOU AGREE TO COMPLY WITH AND TO BE BOUND BY THE TERMS AND CONDITIONS OF THIS AGREEMENT AND ALL APPLICABLE LAWS AND REGULATIONS IN THEIR ENTIRETY WITHOUT LIMITATION OR QUALIFICATION. IF YOU DO NOT AGREE TO BE BOUND BY THIS AGREEMENT, THEN YOU MAY NOT ACCESS OR

☒ By clicking on the box, you indicate that you have read and agree to this [Developer Agreement](#) and the [Twitter Developer Policy](#), additionally as it relates to your display of any of the Content, the [Display Requirements](#); as it relates to your use and display of the Twitter Marks, the [Twitter Brand Assets and Guidelines](#); and as it relates to taking automated actions on your account, the [Automation Rules](#). These documents are available in hardcopy upon request to Twitter.

[Back](#) [Submit](#)

[PRIVACY](#) [COOKIES](#) [TWITTER TERMS & CONDITIONS](#) [DEVELOPER POLICY & TERMS](#) © 2022 TWITTER INC. [FOLLOW @TWITTERDEV](#) [SUBSCRIBE TO DEVELOPER NEWS](#) ▾

# Access Key

- Lastly, generate Access Key and Access Token Key.
- From the side menu, go to Projects & Apps > Project 1.
- Under Apps, click on the key icon.

The screenshot displays the Twitter Developer Portal interface. On the left, a dark sidebar contains the 'Developer Portal' logo and a navigation menu with options: Dashboard, Projects & Apps (highlighted with a red arrow), Project 1, NewUserApp, Products (marked with a 'NEW' badge), and Account. The main content area is titled 'Project 1' and has tabs for 'Overview' (selected) and 'Settings'. Under the 'Overview' tab, there is a 'Usage' section showing 'MONTHLY TWEET CAP USAGE' with a progress bar at 0% and a '0 Tweets pulled of 2,000,000' limit. Below this is an 'Apps' section listing 'NewUserApp' with a 'Manage' link and a key icon (highlighted with a red arrow). On the right, a 'Helpful docs' sidebar lists links for 'About Projects', 'About Apps', 'About authentication', 'About Tweet caps', and 'Authentication best practices'. At the bottom, a footer contains links for Privacy, Cookies, Twitter Terms & Conditions, Developer Policy & Terms, Copyright 2022 Twitter Inc., Follow @TWITTERDEV, and Subscribe to Developer News.

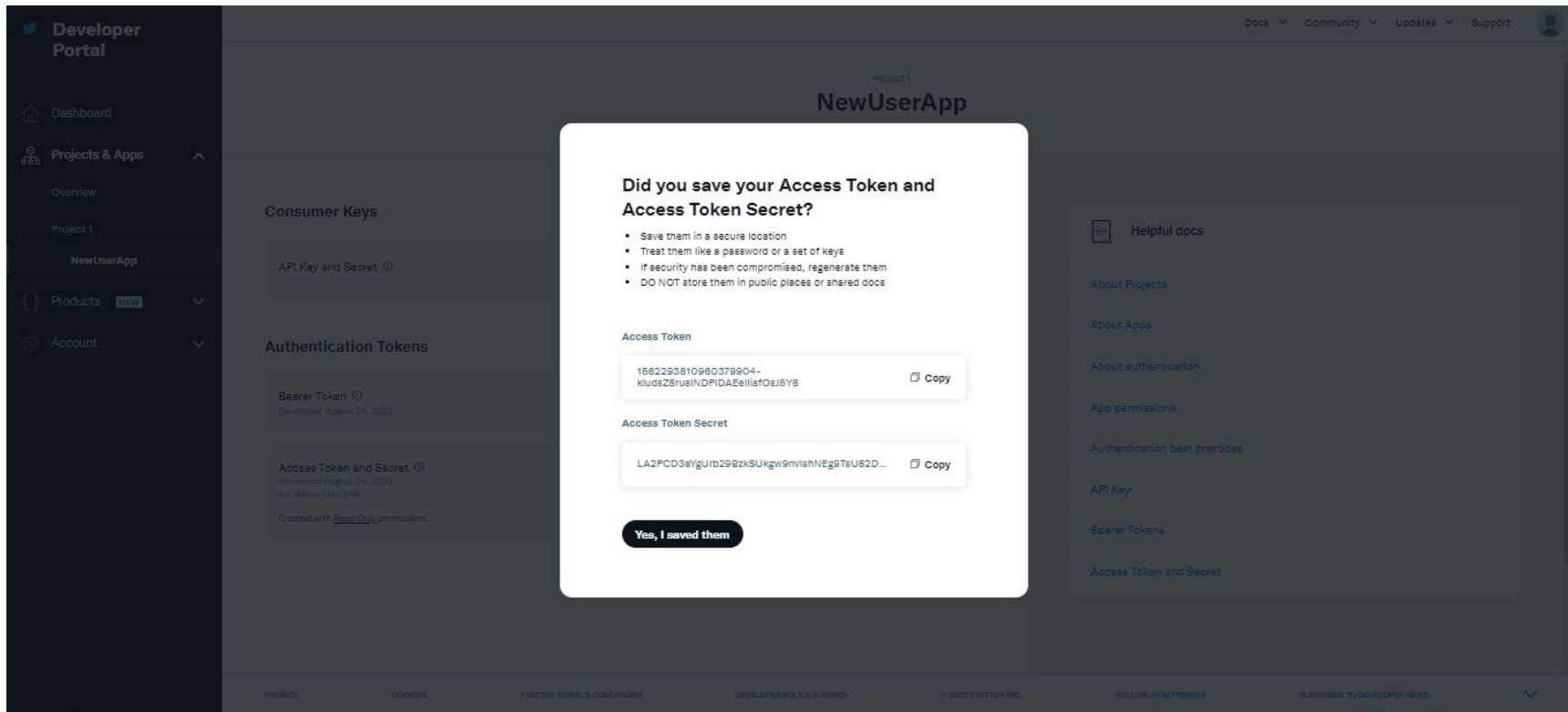
# Access Key

- Click on Generate.

The screenshot shows the Twitter Developer Portal interface. On the left is a dark sidebar with the 'Developer Portal' logo and navigation links: 'Dashboard', 'Projects & Apps' (expanded to show 'Overview' and 'Project 1'), and 'NewUserApp' (selected). Below these are 'Products' (with a 'NEW' badge) and 'Account'. The main content area is titled 'NewUserApp' with tabs for 'Settings' and 'Keys and tokens'. Under 'Consumer Keys', there is a section for 'API Key and Secret' with a 'Reveal API key hint' link and a 'Regenerate' button. Under 'Authentication Tokens', there are two sections: 'Bearer Token' (with a 'Generated August 24, 2022' timestamp, 'Revoke', and 'Regenerate' buttons) and 'Access Token and Secret' (with a 'For @NewUser248' timestamp and a 'Generate' button). A red arrow points to the 'Generate' button. On the right, a 'Helpful docs' sidebar lists links: 'About Projects', 'About Apps', 'About authentication', 'App permissions', 'Authentication best practices', 'API Key', 'Bearer Tokens', and 'Access Token and Secret'. The footer contains links for 'Privacy', 'Cookies', 'Twitter Terms & Conditions', 'Developer Policy & Terms', '© 2022 Twitter Inc.', 'Follow @TwitterDev', and 'Subscribe to Developer News'.

# Access Key

- It will load a popup form with the Access Key and Access Token Key.
- Save them into a notepad/external file as they are required for further analysis and will not be available later.
- Click on Yes, I saved them.



# Authentication of Twitter Account in Python

```
# Install "tweepy" library from Anaconda Prompt
```

```
pip install tweepy
```

```
# Import "tweepy" library in python
```

```
import tweepy
```

- ❑ **tweepy** is a Python library which provides access to the Twitter API.

- Copy the API key and API secret key from slide 11 and Access token and Access token secret from slide 23 and paste in Python code.

```
consumer_key = " paste here your API key "  
consumer_secret = " paste here your API secret key "  
access_token = " paste here your Access token "  
access_token_secret = " paste here your Access token secret "
```

- Complete the twitter authorization process.

```
auth = tweepy.OAuthHandler(consumer_key, consumer_secret)  
auth.set_access_token(access_token, access_token_secret)  
api = tweepy.API(auth)
```

- Twitter account is now connected to Python for fetching tweets.

- ❑ **OAuthHandler()** creates the authentication object.
- ❑ **set\_access\_token()** sets your access token and secret
- ❑ **API()** creates the API object while passing in auth information

# Fetching Data From Twitter

# Fetch tweets for Delhi pollution.

```
tweets = tweepy.Cursor(api.search_tweets ,
                        q='pollution + delhi',
                        lang="en").items(500)
pollution_tweets=[tweet.text for tweet in tweets]
pollution_tweets[0:5]
```

- ☐ **Cursor()** function will issue a search of Twitter based on a supplied search string.
- ☐ First argument search query to issue to twitter. Use "+" to separate query terms.
- ☐ **.items()** specifies the maximum number of tweets to return
- ☐ **lang="en"** search tweets in English language.
- ☐ Twitter developer account searches tweets for maximum 7 days.



**Note :** The tweets fetched will be different depending upon when they are extracted.



# Fetching Data From Twitter

# Output:

```
['RT @AamAadmiParty: "I acknowledge that the high-level of pollution in Delhi is  
mainly due to its local sources... and thus, we are conducting...',  
 'T-1 \nAs the Winds started Blowing in delhi the concern about the pollution also  
get blown away #DelhiPollution #DelhiGovt Shame Shame',  
 '@blkahn Plus Amazon fires, Delhi pollution, China, ocean plastics...must find  
solutions',  
 '@Sharmistha_GK @ArvindKejriwal Pollution is not just a Delhi problem. Its a  
north India problem. If you view the ma... https://t.co/gDHPbiYc2T',  
 'RT @AamAadmiParty: "I acknowledge that the high-level of pollution in Delhi is  
mainly due to its local sources... and thus, we are conducting...']
```

# Cleaning Data For Text Analysis

# Clean this corpus before we make the word cloud.

# Convert all data to lowercase

```
pollution_tweets2 = [item.lower() for item in pollution_tweets]
```

# Remove twitter handles

```
import re
pollution_tweets2 = [re.sub('@[^\s]+','',item) for item in
pollution_tweets2]
```

# Remove twitter hyperlinks

```
pollution_tweets2 = [re.sub('(http\S+)','',item) for item in
pollution_tweets2]
```

# Remove Punctuation

```
from string import punctuation
remove_punc = str.maketrans("", "", punctuation)
pollution_tweets2 = [item.translate(remove_punc) for item in
pollution_tweets2]
```

# Remove Numbers

```
from string import digits
remove_digits = str.maketrans('', '', digits)
pollution_tweets2 = [item.translate(remove_digits) for item in
pollution_tweets2]
```

# Cleaning Data For Text Analysis

# Remove stopwords

```
import nltk
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
additional = ['rt','rts','retweet']
swords = set().union(stopwords.words('english'),additional)
pc=[]
for item in pollution_tweets2:
    word_tokens = word_tokenize(item)
    pol_clean = [w for w in word_tokens if not w in swords]
    pc.append(pol_clean)
pc[2]
```

❑ **set().union(stopwords.words('english'), additional)** creates a set of all the english stopwords which enables us to use that set entirely to remove stop words from the data.

# Remove all white space created due to above cleaning

```
rm_ws=[]
for item in pc:
    remove_whitespace = [x.strip() for x in item]
    rm_ws.append(remove_whitespace)
```

# Remove search words that are more frequent

```
import itertools
combined = list(itertools.chain.from_iterable(rm_ws))
remove_common = [w for w in combined if not w in "pollution+delhi"]
```

# Generating WordCloud

```
# Generate wordcloud of clean data
```

```
from wordcloud import WordCloud
wordcloud =
WordCloud(background_color="white").generate(str(remove_common))
import matplotlib.pyplot as plt
# plot the WordCloud image
plt.figure(figsize = (8, 8), facecolor = None)
plt.imshow(wordcloud); plt.axis("off")
plt.tight_layout(pad = 0); plt.show()
```

## # Output



# Sentiment Analysis Using 'vader'

```
# Import "SentimentIntensityAnalyzer" from "vader"  
# Import "pandas"
```

```
from nltk.sentiment.vader import SentimentIntensityAnalyzer  
import pandas as pd
```

```
# Perform Sentiment analysis
```

```
sent_analysis = pd.DataFrame(columns =  
['sentence', 'compound', 'negative', 'neutral', 'positive'])  
sid = SentimentIntensityAnalyzer()  
for i in range(0, len(pollution_tweets2)):  
    ss = sid.polarity_scores(pollution_tweets2[i])  
    compound = ss['compound']  
    negative = ss['neg']  
    neutral = ss['neu']  
    positive = ss['pos']  
    sent_analysis = sent_analysis.append({"sentence":  
pollution_tweets2[i], "compound": compound, "negative":  
negative, "neutral": neutral, "positive": positive},  
ignore_index=True)  
sent_analysis
```

# Sentiment Analysis Using "vader"

# Output :

sentence	compound	negative	neutral	positive
California, even with the fires, doesn't compare to Singapore's worst day 0and is a tiny fraction of Delhi's best da... <a href="https://t.co/KeogdwDVtG">https://t.co/KeogdwDVtG</a>	0.0258	0.156	0.684	0.16
RT @ndtv: Devotees stand knee-deep in toxic foam in Delhi's Yamuna for 1#ChhathPuja. <a href="https://t.co/tiHnluBdNz">https://t.co/tiHnluBdNz</a> <a href="https://t.co/D8r6sVVpI5">https://t.co/D8r6sVVpI5</a>	0.128	0	0.903	0.097
RT @PopovichN: We visualized microscopic air pollution that wreaks havoc on 2human health. Compare your city's air quality to some of the wo...	-0.5994	0.151	0.849	0
RT @PopovichN: We visualized microscopic air pollution that wreaks havoc on 3human health. Compare your city's air quality to some of the wo...	-0.5994	0.151	0.849	0
RT @KailasK86985883: @BBMP_MAYOR We are not far from Delhi in Air Pollution!! Time is NOW to take right decision. 4Please stop the proposals...	0.1739	0.084	0.805	0.111
RT @nytcclimate: We visualized the damaging, tiny particles that wreak havoc 5on human health. From the Bay Area to New Delhi, see how the wo...	-0.802	0.238	0.762	0
RT @nytcclimate: We visualized the damaging, tiny particles that wreak havoc 6on human health. From the Bay Area to New Delhi, see how the wo...	-0.802	0.238	0.762	0
RT @nytcclimate: We visualized the damaging, tiny particles that wreak havoc 7on human health. From the Bay Area to New Delhi, see how the wo...	-0.802	0.238	0.762	0
RT @EuroGeosciences: See how the world's most polluted air compares with your city's: 8@nytcclimate has visualized the damaging, tiny partic...	-0.7645	0.268	0.732	0
RT @nytimes: Particulate pollution in the air soared last month in New 9Delhi. But the city struggles with air quality throughout the year:...	-0.3612	0.102	0.898	0
RT @ismaelnafria: Gran interactivo del @nytimes. Perfecto ejemplo de info 10útil y personalizada - See How the World's Most Polluted Air Comp...	-0.2484	0.14	0.763	0.097

## Interpretation:

- Negative compound score indicates, negative sentiments. Compound score give sum of all the lexicons standardized between -1 and 1.

# Quick Recap

## Twitter Data

- Twitter data constitutes a rich source that can be used for capturing information about any topic.

## Libraries in Python

- Install libraries '**tweepy**' and '**nltk**'

## Steps required before fetching data from Twitter

- Create twitter account.
- Create developer account.
- Create app.
- Authentication of Twitter account in Python

## Function in Python

- **tweepy.Cursor()** function extracts data from Twitter.