



# **STREAMING TWITTER DATA INTO MySQL DATABASE**

## **PROJECT REPORT**

### **GROUP MEMBERS:**

- 1. SATWIK TANWAR (20BCE2945)**
- 2. ABHAY NIMMAGADDA (20BCE2327)**
- 3. SHIVA RAMAKRISHNA Y (20BCE2864)**
- 4. DHYANA SAI N (20BCE2956)**
- 5. VARANASI PHANI BALA MOUNIKA (20BCE2963)**

**SUBMITTED TO:** Prof. Jyotismita Chaki

**COURSE NAME:** Database Management Systems

**COURSE CODE:** CSE2004

**SLOT:** L57+L58

## **ABSTRACT**

It's undeniable that social media has now become an extremely important platform to aid businesses to expand their online reach. Marketers have gradually become increasingly conscious of the relevance of social media platforms such as Facebook, Twitter, and LinkedIn in business growth via word-of-mouth marketing over the years. The main aim of our project is to stream twitter data into MySQL database. We can stream data directly into a MySQL database using Python and an API. We'll use the Twitter API to look for tweets and users we are interested in and import them directly into our database. The data will be available for further examination at any moment once we've completed this.

## **Table of Contents:**

### **1. Introduction**

- 1.1. How did the project evolve?**
- 1.2. Why is this helpful?**

### **2. Background**

- 2.1. Previous works**
- 2.2. Drawbacks of Previous Designs**

### **3. Methodology**

- 3.1. Full Description of Project**
- 3.2. Full Architecture**
- 3.3. Obtained Results**

### **4. Conclusion**

### **5. References**

## **1. INTRODUCTION:**

### **1.1. How did this project evolve?**

If we recollect the purpose of social media, we can say that it is to bring people together from all over the world. It's understandable why businesses have rushed to social media, since they've realised it's a terrific location for communication, whether it's good news, terrible news, fake news, or an international disaster. As a result, it would be preferable if we could store data from Twitter in a MySQL database.

### **1.2. Why is this helpful?**

We'll use the Twitter API to get data about the followers of an account and feed it directly into our database. This data can be used to analyse their audience and increase their reach. We will store data like the username of the person who posted the tweet, the time it was posted, the tweet, the number of retweets it received, number of likes and so on. After that, the data will be available for analysis at any time. Companies will benefit greatly from this since they will be able to see how people are reacting.

## **2. BACKGROUND**

### **2.1. Previous Works**

#### a) MOA-TweetReader: Real-Time Analysis in Twitter Streaming Data

- Author: Albert Bifet, Geoffery Holmes, and Bernhard Pfahringer University Of Waikato, Hamilton, New Zealand
- Tweets are preprocesssed and converted MOA-TweetReader to vectors of attributes or machine learning instances. The second component of the system is a learner trained with several instances, that is able to

- predict the class label of incoming unlabeled instances. Finally, a change detector monitors the predictions, and outputs an alarm signal when change is detected.
- Application: detecting changes in term frequencies and performing real time sentiment analysis.

b) Automating Corpora Generation with Semantic Cleaning and Tagging of Tweets for Multi-dimensional Social Media Analytics

- Author: Nazura Javed Research Scholar Bangalore University Bangalore, India and Muralidhara B.L. Bangalore University Bangalore, India.
- The ‘processTweet’ algorithm which automates the process of collecting and cleaning the tweets was coded using Python.
- Unprocessed as well as the cleaned, processed tweets were stored in MySQL database.
- PyEnchant, a spell checking library for Python was used for checking the spelling of English words and for splitting the joined words.
- Jellyfish, a Python library, was used for doing approximate and phonetic matching using the metaphone algorithm. A customized slang/abbreviation dictionary was used for replacing the standard slangs and abbreviations. The experiment was conducted by filtering English tweets relating to Indian politics and Prime Minister ‘Narendra Modi’.

c) Twitter Data Analysis Using Spark

- Author: Yash Bopardikar, California State University, Sacramento
- Set spark worker instances to 4, which initiates 4 worker instances as soon as the spark cluster is started. The Executor instances are set to 5 and each worker is allocated 8 GB of memory i.e. the spark cluster can use the total amount of memory available on the system. The spark Executor and Worker cores are set to 6 and 10 which are virtual cores.
- Use twitter4j library which is embedded into TwitterUtils library for getting specific fields out of the data.

- Sentiment analysis is performed using Stanford's Natural Language Processing library. It takes text input and sends this library to get the sentiment in return.

d) Public Perception of the COVID-19 Pandemic on Twitter: Sentiment Analysis and Topic Modelling Study

- Author: Sakun Boon-Itt, Yukolpat Shunkan (Originally published in JMIR Public Health and Surveillance)
- This study was conducted in two stages: (1) data collection using the Twitter streaming API to collect COVID-19 related posts in the English Language, and (2) data analysis to identify trends, keywords, and themes.
- The objective of this study was to answer questions related to themes, public concerns, and sentiments regarding the COVID-19 pandemic through social media analytics.
- The tweet database was created by specifying keywords and metadata such as language, source, data range, and location. This search used keywords and specific hashtags (#) such as *coronavirus*, *covid\_19*, *2019-nCov*, and *covid-19* in the English language using the *searchtweets* tool. The tweets were then converted into a corpus.

## 2.2. Drawbacks of Previous Designs

- It is difficult to maintain the structure of data without using a DBMS.
- There is more redundancy in previous works due to use of only scripting language.
- There is no constraint as to what kind of tweets we have to extract which may increase variance while using data for machine learning algorithms.

### **3. METHODOLOGY:**

#### **3.1. Full Description of the Project**

- Twitter is a micro blogging service that allows people to communicate with short 140 character messages. Twitter has become a platform for information dissemination and change. Twitter provides an open platform for expression, and hence mining of tweets can provide valuable insight into political scenario and societal opinion. Tweets can be collected using the REST or Streaming API.
- Although the tweet text is limited to just 140 characters, there is an associated rich metadata like the location, time zone, geographical latitude, geographical longitude, creation date, followers count and retweet count that can be captured and utilized for analysis. The retweet count is an indicator of the popularity of the tweet. The geographical parameters and location indicate the geographical origin of the tweet.
- Mining the text in conjunction with metadata can provide an insight into the different political perspectives, political trends, sentiments and opinions.
- It is required that the user has to have certain things such as a twitter account and API credentials, access to mysql database and libraries such as tweepy and mysql connector.
- We created a configuration code to extract keys such as API\_Key, Access token, secret key etc.
- Then we have written an another function to essentially link the database with the details we have extracted from the previous code and create a table to store the tweets which are to be streamed into the database.
- We are using the function createTableUser to store the followers or users in general in a table with user ID as primary key, name, username, description text, account created at, location the tweet was tweeted, followers count, following people count, tweet count and even the profile image URL as the attributes.
- We have also created another function createTavleTweets to store the tweets in a table with tweetID as the primary key, userID as the foreign key, authorID, account created at, tweet text, retweet count, reply count, like count, quote count as the attributes.
- We have also used functions like filterOnFollowers to get the number of active users, ifTweetPresent to chech whether the tweet is present in our

database or not, insertTweet to insert data in tweets table, insertFollower to insert data in users table.

- We are going to use insertInDB function to retrieve the followers of the input username and then creating a table to store the details of those users.
- We are also using the function insertFilteredTweets to retrieve only the filtered tweets which pass the conditions like there have to be minimum of 1000 followers of the person who is following the userID we gave as input, and minimum tweet count as 10. We are also gathering the general public reaction to their tweets by storing the count of likes the tweet got, the number of times it was retweeted etc.
- Then we can perform whatever operation we want to execute on the dataset we collected.
- For the front-end part we have used basic HTML, CSS and flask for visual representation purposes.

## 3.2. Full Architecture

### Code:

#### 1. config.py

```
class Config(object):
    DEBUG = True
    TESTING = False

    class DevelopmentConfig(Config):
        SECRET_KEY = "12345"

        config = {
            'development': DevelopmentConfig,
            'testing': DevelopmentConfig,
            'production': DevelopmentConfig
        }

        #API Config
        API_KEY = 'API key'
        API_KEY_SECRET = 'API_KEY_SECRET'
        BEARER_TOKEN = 'BEARER_TOKEN'
        ACCESS_TOKEN = 'Access Token'
        ACCESS_TOKEN_SECRET = 'Access Token Secret'

    #Database Config
    host='localhost'
```

```
user='root'  
password='password'  
database='TwitterDB'
```

## 2. mySqlConnection.py

```
from colorama import Cursor  
import mysql.connector  
  
#creating connector object to connect with database  
def connect(host, user, password, database):  
  
    conn = mysql.connector.connect(  
        host=host,  
        user=user,  
        password=password,  
        database=database  
    )  
    return conn  
  
#creating table to store followers or user in general  
def createTableUser(conn,tableName):  
    cursor=conn.cursor()  
    sql = "create table "+ str(tableName)+"""(  
        userID bigint primary key,  
        name varchar(100),  
        username varchar(100),  
        description text,  
        created_at varchar(45),  
        location varchar(100),  
        followers_count int,  
        following_count int,  
        tweet_count int,  
        verified boolean,  
        profileImageURL varchar(255)  
    );  
    """  
    cursor.execute(sql)  
    cursor.close()  
  
#creating table to store tweets  
def createTableTweets(conn,tableName,userTableName):  
    cursor=conn.cursor()  
    sql = "create table "+ str(tableName)+"""(  
        tweetID bigint,  
        userID bigint,  
        authorID bigint,  
        created_at varchar(45),  
        tweet text,  
        retweet_count int,  
        reply_count int,  
        like_count int,  
        quote_count int,  
        PRIMARY KEY(tweetID),
```

```

        FOREIGN KEY(userID) REFERENCES """
+userTableName+" (userid) ;"

cursor.execute(sql)
cursor.close()

#inserting data in users tables
def insertFollower(conn, recordFollower, tableName):
    cursor = conn.cursor()
    query="INSERT INTO "+ tableName +"""
        (userID, name, username, description, created_at,
location, followers_count, following_count, tweet_count,
verified, profileImageURL)
        VALUES (%s,%s,%s,%s,%s,%s,%s,%s,%s,%s) """
    cursor.execute(query,recordFollower)
    conn.commit()
    cursor.close()
    #inserting data in tweets table
def insertTweet(conn, recordTweet,tableName):
    cursor = conn.cursor()
    query="INSERT INTO " + tableName +"""
        (tweetID, userID, authorID, created_at, tweet,
retweet_count, reply_count, like_count, quote_count)
        VALUES (%s,%s,%s,%s,%s,%s,%s,%s,%s) """
    cursor.execute(query,recordTweet)
    conn.commit()
    cursor.close()
    #query to get active users
def filterOnFollowers(conn,tableName,minFollowers,minTweetCount):
    cursor=conn.cursor()
    query=("SELECT username FROM " +tableName+ " WHERE
followers_count > "+str(minFollowers) + " and tweet_count >
"+str(minTweetCount))
    cursor.execute(query)
    result = cursor.fetchall()
    cursor.close()
    return result

#function to check if a tweet is present
def ifTweetPresent(conn, tableName, tweetID):
    cursor=conn.cursor()
    query = ("SELECT tweetID from "+tableName+" where tweetID =
") +str(tweetID)
    cursor.execute(query)
    result=cursor.fetchall()
    cursor.close()
    if len(result)!=0: return True
    else: return False

```

### 3. search.py

```
#imports
import tweepy
import config
import mysqlConnection

#setting up a client to make API request to twitter API
def createClient():
    client = tweepy.Client(bearer_token=config.BEARER_TOKEN,
                           consumer_key=config.API_KEY,
                           consumer_secret=config.API_KEY_SECRET,
                           access_token=config.ACCESS_TOKEN,
                           access_token_secret=config.ACCESS_TOKEN_SECRET)
    return client

#Retriving followers of the input username and then creating a table to
#store the details of those users
def insertInDB(client,yourUsername):

    #getting id of user from username
    id=client.get_user(username=yourUsername).data.id

    #fields that we want to extract

    user_fields=['description','created_at','location','public_metrics','verified','profile_image_url']

    #getting the followers of the input user

    followers=client.get_users_followers(id,user_fields=user_fields,max_results=1000)

    #forming a table to store the data
    conn=mysqlConnection.connect(config.host, config.user,
                                config.password, config.database) #mysql connector object

    #creating the table
    tableName=yourUsername+"Followers"
    mysqlConnection.createTableUser(conn,tableName)

    #inserting data in table
    for user in followers.data:

        record=(user.id,user.name,user.username,user.description,user.created_at,
                user.location,user.public_metrics['followers_count'],user.public_metrics[
                    'following_count'],user.public_metrics['tweet_count'],user.verified,user.
                profile_image_url)

        mysqlConnection.insertFollower(conn,record,tableName)

    #getting posted and liked tweets of active users
```

```

#active user are based upon minimum followers they have and the total
number of tweets they've done.

def
insertFilteredTweets(client,userName,minFollowers=1000,minTweetCount=10):

    conn = mysqlConnection.connect(config.host, config.user,
config.password, config.database) #mysql connection object

    #getting usernames of active users
    followerTableName=userName+'Followers'

    userNames=mysqlConnection.filterOnFollowers(conn,followerTableName,minFol
lowers,minTweetCount)

    #getting tweets
    for user in userNames:
        tweet_fields=['author_id','created_at','geo','public_metrics']

        #getting tweets posted by the user
        query='from:'+user[0]+' -is:retweet'
        tweetsPosted =
client.search_recent_tweets(query=query,tweet_fields=tweet_fields,
max_results=100)

        #getting tweets liked by the user
        userID=client.get_user(username=user[0]).data.id
        tweetsLiked = client.get_liked_tweets(id=userID,
tweet_fields=tweet_fields)

        if tweetsPosted.data==None or tweetsLiked.data==None: continue

        #creating table for the user
        userTweetTable = user[0] + "_" + userName + "_Tweets"

mysqlConnection.createTableTweets(conn,userTweetTable,followerTableName)

    #function to insert tweets
    def insert(tweets,tableName(userID):
        for tweet in tweets.data:

            record=(tweet.id,userID,tweet.author_id,tweet.created_at,tweet.text,tweet
.public_metrics['retweet_count'],tweet.public_metrics['reply_count'],twee
t.public_metrics['like_count'],tweet.public_metrics['quote_count'])

            #checking if tweet already inserted....
            if
mysqlConnection.ifTweetPresent(conn,tableName,tweet.id): continue

            #inserting the tweet
            mysqlConnection.insertTweet(conn,record,tableName)

            insert(tweetsPosted,userTweetTable,userID)
            insert(tweetsLiked,userTweetTable,userID)

```

## 4. app.py

```
from flask import Flask, render_template, url_for, redirect, request
import config
import search

def page_not_found(e):
    return render_template('404.html'), 404

app = Flask(__name__)
app.config.from_object(config.config['development'])

app.register_error_handler(404, page_not_found)

@app.route('/', methods=["GET", "POST"])
def index():
    client=search.createClient()
    if request.method == 'POST':
        userName=request.form['username']
        minFollowers=request.form['minFollowers']
        minTweetCount = request.form['minTweetCount']

        search.insertInDB(client,userName)

    search.insertFilteredTweets(client,userName,minFollowers,minTweetCount)

    return render_template('index.html', **locals())

@app.route('/retweet/<string:tweetId>', methods=["GET", "POST"])
def reweet(tweetId):

    return redirect(url_for('index'))

if __name__ == '__main__':
    app.run(host='0.0.0.0', port='8000', debug=True)
```

## 5. index.html

```
<!doctype html>
<html lang="en">
  <head>
    <!-- Required meta tags -->
    <meta charset="utf-8">
    <meta name="viewport" content="width=device-width,
initial-scale=1">

    <title>Twitter Followers Analysis</title>
    <link rel="shortcut icon" type="image/x-icon" href="{{
url_for('static', filename='images/favicon.png') }}">

    <!-- Bootstrap CSS -->
    <link
      href="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/css/bootstrap.min.css" rel="stylesheet"
      integrity="sha384-1BmE4kWBq78iYhFdvKuhfTAU6auU8tT94WrHftjDbrCEXS
      U1oBoqyl2QvZ6jIW3" crossorigin="anonymous">

  </head>
  <body>

    <div class="container mt-5 mb-5">

      <h1 class="mb-3">Twitter Follower Analysis</h1>

      <form class="" action="/" method="post">

        <div class="col-sm-3">
          <label class="sr-only">Username</label>
          <div class="input-group-prepend">
            <div class="input-group-text">@</div>
          </div>
          <input type="text" class="form-control"
placeholder="Username" id="username" name="username">
        </div>
      
```

```

<div class="mb-3">
    <br><br>
    <h3 class="mb-3">Apply filters to retrive posted and
liked tweets of your active followers.</h3>
    <label for="minFollowers" class="form-label">Minimum
Followers Count</label>
    <input type="text" class="form-control"
id="minFollowers" name="minFollowers" placeholder="1000....">
</div>

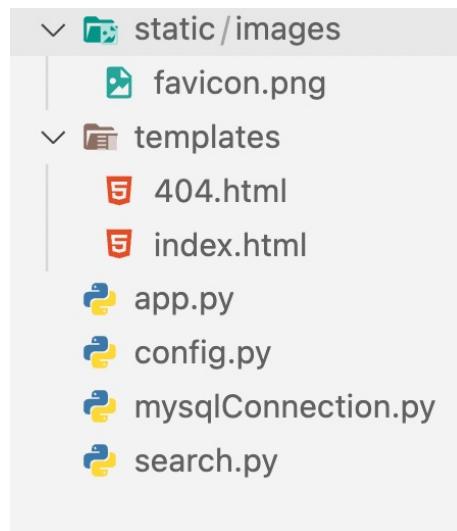
<div class="mb-3">
    <label for="minTweetCount" class="form-label">Minimum
Tweets Count</label>
    <input type="text" class="form-control"
id="minTweetCount" name="minTweetCount" placeholder="10....">
</div>

<button type="submit" class="btn btn-primary"
onclick="result()"> Find Followers </button>
<p id="output"></p>
<br>
<br>
<script> function result() {
    document.getElementById("output").innerHTML = "Data has
been imported to database";
}
</script>
</form>
</div>

<!-- Option 1: Bootstrap Bundle with Popper -->
<script
src="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/js/bootstrap.bundle.min.js"
integrity="sha384-ka7Sk0Gln4gmtz2MlQnikT1wXgYsOg+OMhuP+IlRH9sENBO
0LRn5q+8nbTov4+lpx crossorigin="anonymous"></script>
</body>
</html>

```

## File Structure:



## Output:

### Web Interface

Not Secure | 172.17.40.80:8000

### Twitter Follower Analysis

Username  
@  
elonmusk

Apply filters to retrieve posted and liked tweets of your active followers.

Minimum Followers Count  
1000

Minimum Tweets Count  
10

**Find Followers**

Data has been imported to database

## Tables:

TwitterDB	
Tables	
>	AAAALMARRI_elonmusk_Tweets
>	AlfredMwandagha_elonmusk_Twe...
>	alixestriker_elonmusk_Tweets
>	amodi_84_elonmusk_Tweets
>	A_Rov_elonmusk_Tweets
>	Asadchaudhryusa_elonmusk_Twe...
>	BedfordBryce_elonmusk_Tweets
>	BinhaRio2019_elonmusk_Tweets
>	bsumanbhatarai_elonmusk_Tweets
>	danish_swaleh_elonmusk_Tweets
>	demonryuzaki13_elonmusk_Tweets
>	elonmuskFollowers
>	endrum_ak_elonmusk_Tweets
>	globaltravel62_elonmusk_Tweets
>	grifindor_elonmusk_Tweets
>	gurgel_maria_elonmusk_Tweets
>	HuntTaz_elonmusk_Tweets
>	JaeTizzle_elonmusk_Tweets
>	JoeBrunoWSOC9_elonmusk_Twe...
>	LakewoodPDCO_elonmusk_Tweets
>	Legalparrot71_elonmusk_Tweets
>	ManvendraPandav_elonmusk_Twe...
>	MarionMarechal_elonmusk_Tweets
>	maytevillares_elonmusk_Tweets
>	mjsamman_elonmusk_Tweets
>	OscarG_1978_elonmusk_Tweets
>	pablishSLP_elonmusk_Tweets
>	recomtogadaraya_elonmusk_Twe...
>	R_KuehnAfD_elonmusk_Tweets

# Followers Table Structure:

Local Instance 3306

Query 4 | elonmuskFollowers | Limit to 1000 rows

1 • SELECT \* FROM TwitterDB.elonmuskFollowers;

Result Grid | Filter Rows: Search | Edit: | Export/Import: | Fetch rows: | Result Grid | Form Editor | Field Types | Query Stats | Execution Plan

userID	name	username	description	created_at	location	followers_count	following_count	tweet_count	verified	profileImageURL
6103502	Nishant Modak	nishantmodak	Building @last9Inc . Startups, Tea, B...	2007-05-17 11:45:39	NULL	2714	512	6188	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
6208662	melissa	missy6579		2007-05-21 21:01:50	NULL	69	212	5	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
16110147	Wendi Ellison	poochiebeer75		2008-09-03 08:55:03	Bridgeport, CT	16	189	35	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
16240312	Larkin Kennedy	larkinkennedy		2008-09-11 13:50:42	Mississippi	124	241	333	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
16251168	ZombieMike	MikeyMc	Community Advocate for Castlevania...	2008-09-12 01:43:33	Jacksonville, AR	208	327	3097	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
18863852	sotoksyros	sotoksyros	was a digital worker	2009-01-11 11:06:52	Greece	699	651	60942	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
19342090	Tyler Atkinson	TylerAtkinson	Nobody Puts Bagz In a Corner	2009-01-22 15:14:50	Nashville, TN	38	529	1720	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
20821371	CC	cartercl		2009-02-14 00:37:05	NULL	1	42	0	0	<a href="https://abs.twimg.com/sticky/default_header_normal.png">https://abs.twimg.com/sticky/default_header_normal.png</a>
23141900	Kranyakoot	Kranyakoot		2009-03-07 00:25:24	NULL	16	105	186	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
23664244	Brooke Smith	bsmith14		2009-03-10 21:46:28	Missouri	82	781	337	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
28301028	Bill Bombeck	bomber5550	Husband, Social Studies Teacher, Fo...	2009-03-11 18:05:54	Grand Island, NE	435	1157	1856	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
28149150	Carlos Santos	grifindor	🌐⚡️ Oficial EB RR, Bolsonarista...	2009-03-24 01:50:02	Recife, Brasil	1494	1947	22057	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
26202200	Sadouq Rizvi	SadouqRizvi	Managing Partner @ Top Foods, Lov...	2009-03-24 08:50:26	Mumbai	545	603	7651	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
27830755	Arghyaz Roych	arghyaz		2009-03-31 07:04:11	Kolkata	211	211	1497	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
29046404	Car	carcar_m		2009-04-05 19:40:27	NULL	0	130	0	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
33113761	Josué Junior	josuejunior	Penso, logo existo	2009-04-19 03:22:34	Sao Paulo, Br...	120	163	654	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
33662409	Lauren Mandel	laurevmandel	Still dreaming of being a Radio City...	2009-04-20 21:03:37	San Antonio, TX	215	543	4412	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
35018860	Pablo Rodriguez	publishSLP	Ing. en Electrónica, me hubiera gusta...	2009-04-24 19:47:29	NULL	1529	809	3351	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
35270595	Haldun Demir	haldund		2009-04-25 18:05:35	NULL	125	144	514	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
35474453	JC Chavez	zacksogen	Ad astra per aspera - By hardwork a...	2009-04-26 14:27:27	Philippines	75	207	1554	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
36605917	matthew	matthewd83		2009-04-20 03:37:38	NULL	49	279	389	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
38284833	Tim Westcott	Tim_Westcott		2009-05-06 21:45:46	NULL	9	33	7	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
41749441	Eliana	elianavan		2009-05-22 04:15:06	Peru	103	351	508	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
41851423	H.Khan	mayamains...		2009-05-22 16:51:21	Karachi	306	403	1699	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
43332560	Joe Bruno	JoeBrunoW	Passionate government and breaking...	2009-05-29 13:50:02	Charlotte, NC	40670	3908	88601	1	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
43959383	PrastoMX	PrastoMX	Norteño I Videojuegos I Kokun I Nint...	2009-06-01 20:04:20	Distrito Feder...	369	310	12361	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>
46883071	John Bui	johnbui1994	Rutgers University '16	2009-06-13 13:12:45	New Jersey	187	562	5111	0	<a href="https://pbs.twimg.com/profile_image">https://pbs.twimg.com/profile_image</a>

elonmuskFollowers 1 | Apply | Revert

Query Completed

# Tweets Table Structure:

Local Instance 3306

Query 4 | elonmuskFollowers | alixestriker\_elonmusk\_Tweets | Limit to 1000 rows

1 • SELECT \* FROM TwitterDB.alixestriker\_elonmusk\_Tweets;

Result Grid | Filter Rows: Search | Edit: | Export/Import: | Result Grid | Form Editor | Field Types | Query Stats | Execution Plan

tweetID	userID	authorID	created_at	tweet	retweet_count	reply_count	like_count	quote_count	quote_user_id
151642715692613636	102927265	102927265	2022-04-19 15:44:52	@AugustoAraiza <a href="https://t.co/ny4x5gb...">https://t.co/ny4x5gb...</a>	0	0	0	0	
1516451034750758918	102927265	102927265	2022-04-19 16:17:55	Hoy en el día de la bicicleta, acuérde...	7	16	136	2	
1516454414860574724	102927265	102927265	2022-04-19 16:31:21	@Soyunservil Por nadá!	0	0	1	0	
1516459051764310016	102927265	102927265	2022-04-19 16:49:47	@abraham_CT 🐸	0	1	0	0	
1516466491276857350	102927265	102927265	2022-04-19 17:19:20	@AstroautoCarlo <a href="https://t.co/7PZV...">https://t.co/7PZV...</a>	0	0	1	0	
1516466698521612289	102927265	102927265	2022-04-19 17:20:10	El CrossFit mi pasión. <a href="https://t.co/LE...">https://t.co/LE...</a>	1	1	20	0	
151646748781244678	102927265	102927265	2022-04-19 17:23:18	@TuitierDeCulto Llévame a la playa...	0	1	1	0	
1516468983663599107	102927265	102927265	2022-04-19 17:29:15	@TuitierDeCulto <a href="https://t.co/BHZEZ...">https://t.co/BHZEZ...</a>	0	1	1	0	
1516469245252677650	102927265	102927265	2022-04-19 17:30:17	@poke107 <a href="https://t.co/08KfAdLy">https://t.co/08KfAdLy</a>	0	1	1	0	
1516515074109751304	102927265	102927265	2022-04-19 20:32:23	Me voy hacer la tanga en la playa...	3	11	82	0	
1516516145213415425	102927265	102927265	2022-04-20 00:36:39	@poke107 <a href="https://t.co/bNQze2cLe">https://t.co/bNQze2cLe</a>	0	1	1	0	
1516634990700000333	102927265	102927265	2022-04-20 04:28:54	@DelaRo_ <a href="https://t.co/geI82RiKM">https://t.co/geI82RiKM</a>	0	0	1	0	
1516822696901980166	102927265	102927265	2022-04-20 16:54:46	Ya no sé qué twítear, creo que ha lle...	1	8	43	0	
1516840633897558016	102927265	102927265	2022-04-20 18:06:03	@caiman_leches 🦈	0	0	1	0	
1516841685195628550	102927265	102927265	2022-04-20 18:10:13	Si a usted le dieran a escoger un sup...	1	5	7	1	
1516844263400685568	102927265	102927265	2022-04-20 18:20:28	@godani502 <a href="https://t.co/J4jnpeCdYB">https://t.co/J4jnpeCdYB</a>	0	1	0	0	
1516848357423124480	102927265	102927265	2022-04-20 18:36:44	J.Rihanna no sé quiere o por qué atr...	1	5	20	0	
151685087679475793	102927265	102927265	2022-04-20 18:46:45	Lo vi en SHEIN y solo pensé en lo in...	0	8	34	2	
1516851005958279168	102927265	102927265	2022-04-20 18:47:16	@CarlitoSalinas Viejo borracho	0	1	1	0	
1516856543513387012	102927265	102927265	2022-04-20 19:09:16	@poke107 <a href="https://t.co/zLn8j4PHR">https://t.co/zLn8j4PHR</a>	0	1	1	0	
1516860344249036806	102927265	102927265	2022-04-20 19:24:22	@ConguitoBB <a href="https://t.co/r4hpAh1R">https://t.co/r4hpAh1R</a>	0	0	1	0	
1516891653449129990	102927265	102927265	2022-04-20 21:28:47	Se imaginan pagar \$299 en Netflix n...	6	17	96	3	
1516891830230654977	102927265	102927265	2022-04-20 21:29:29	@LennyKravitz Show the dick, father!	0	1	5	0	
1516893091554332672	102927265	102927265	2022-04-20 21:34:30	@Disordenado Si los hay	0	0	0	0	
151689379404996080	102927265	102927265	2022-04-20 21:37:17	@Robnilo Me parece una ganga	0	0	2	0	
1516895502146617346	102927265	102927265	2022-04-20 21:44:04	Nos subieron el Apple Music casi dos...	1	4	31	0	
1516898251944996864	102927265	102927265	2022-04-20 21:55:00	@ConguitoBB Pedi que me pagaran...	0	1	0	0	
1516898341317185536	102927265	102927265	2022-04-20 21:55:21	@MisterChilano @LennyKravitz No	0	0	1	0	

alixestriker\_elonmusk\_Tweets 1 | Apply | Revert

Query Completed

## **RESULT ANALYSIS:**

- Two types of tables are obtained, one for users(followers) and another for tweets.
- First, we retrieve a relation containing information about the followers of the input user.
- Then we retrieve the tweets posted and liked by the active followers of the input user.
- This data is extremely valuable for a social media manager because they can analyse it to form the best possible strategy for the social media handles of the given company.

## **CONCLUSION:**

We used the Twitter API to look for tweets posted and liked by the followers of the concerned account and imported them directly into our database. We used the Tweepy library in python which made it very simple to connect to the API and begin streaming data. This has demonstrated that setting up a data pipeline using MYSQL is not difficult, especially when using the great modules that python has to offer.

## **REFERENCES AND CITATIONS:**

### **Previous Works:**

1. [https://link.springer.com/chapter/10.1007/978-3-642-24477-3\\_7](https://link.springer.com/chapter/10.1007/978-3-642-24477-3_7)  
Albert Bifet, Geoffrey Holmes, Bernhard Pfahringer ‘MOA-Tweet Reader Real-Time Analysis in Twitter Streaming Data’
2. <https://publichealth.jmir.org/2020/4/e21978>  
Sakun Boon-Itt, Yukolpat Shunkan (Originally published in JMIR Public Health and Surveillance) ‘Public Perception of the COVID-19 Pandemic on Twitter: Sentiment Analysis and Topic Modeling Study’
3. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.735.944&rep=rep1&type=pdf>

Nazura Javed Research Scholar Bangalore University Bangalore, India  
and Muralidhara B.L. Bangalore University Bangalore, India

‘Automating Corpora Generation with Semantic Cleaning and Tagging  
of Tweets for Multi-dimensional Social Media Analytics’

4. <https://scholarworks.calstate.edu/downloads/db78tc01m>

Yash Bopardikar, California State University, Sacramento ‘Twitter  
Data Analysis Using Spark’

### **Documentation:**

1. <https://docs.tweepy.org/en/stable/client.html>
2. <https://dev.mysql.com/doc/connector-python/en/>