

Comparative Analysis of Users on Twitter to Assess Some Indicators of Social Tension

PROJECT REPORT

Submitted By:

Salonee Gupta	16BCE0194
Japneet Kaur Saluja	16BCE0662
Rishab Gupta	16BCE0757

Course Code: CSE 4022

Natural Language Processing

Professor Arivoli

School of Computer Science and Engineering



VIT[®]
Vellore Institute of Technology

October 2019

CERTIFICATE

This is to certify that the project work entitled “**Comparative Analysis of Users on Twitter to Assess Some Indicators of Social Tension**” that is being submitted by **Salonee Gupta** and **Japneet Kaur Saluja** and **Rishab Gupta** for the course CSE4022 - “Natural Language Processing” is a record of bonafide work done under my supervision. The contents of this project work in full or in parts, have neither been taken from any other source nor have been submitted for any other course.

Signature of Faculty

ACKNOWLEDGEMENT

We would like to thank our professor **Arivoli A.** for her precious guidance and constant support and the Dean of School of Computer Science and Engineering for their pleaded information and opportunity given to us for completion of our project.

Submitted by:

Table of Contents

S. No	Topic	Page #
1.	Abstract	5
2.	Introduction	6
3.	Methodology	7
4.	Tracking	9
	3.1 Keyword Related	9
	3.2 Actor Related	11
5.	Preparation	12
6.	Analysis	15
	5.1 Social Network Analysis	15
	5.2 Sentiment Analysis	18
	5.3 Content Analysis	22
7.	Conclusion	29
8.	References	35
9.	Appendices	36

ABSTRACT

The social media platform Twitter has been used by individuals to communicate on various issues and trends, including social movements, politics and pop culture. This report is aimed at understanding how Twitter users respond to the world around them, and to take inferences of the importance of major issues, their relevance, percentage of positive, negative and unbiased opinion, and their influence on communities.

Our objective is to come up with a mathematical algorithm to answer these questions, and come up with interesting analysis on today's twitter users.

INTRODUCTION

Sentiment analysis – otherwise known as opinion mining – is a much bandied about but often misunderstood term. In essence, it is the process of determining the emotional tone behind a series of words, used to gain an understanding of the attitudes, opinions and emotions expressed within an online mention.

Sentiment analysis is extremely useful in social media monitoring as it allows us to gain an overview of the wider public opinion behind certain topics. The concept of "social tension" expresses a certain state of social groups, their dissatisfaction with the existing situation, the attitude towards what is happening and to other groups. Social dissatisfaction is the dominant factor in determining the level of social tension, therefore, the assessment of social tension is usually based on determining the measure of respondents' satisfaction with the various aspects of their life activity and the conditions of the external social environment. This provision is based on the basic principles of the theory of deprivation, according to which, the increase in dissatisfaction (subjectively-estimated individual gap between the actual and desired life situation) is accompanied by an increase in preparedness for action, including destructive.

The government has not yet formed adequate mechanisms for feedback from social systems. In this regard, the inability to use the institutional system of broadcasting to the government the causes of social tension lead to the fact that protest activity increases and closes its output at the regional level, realizing in various forms of social, political, economic and psychological deviation.

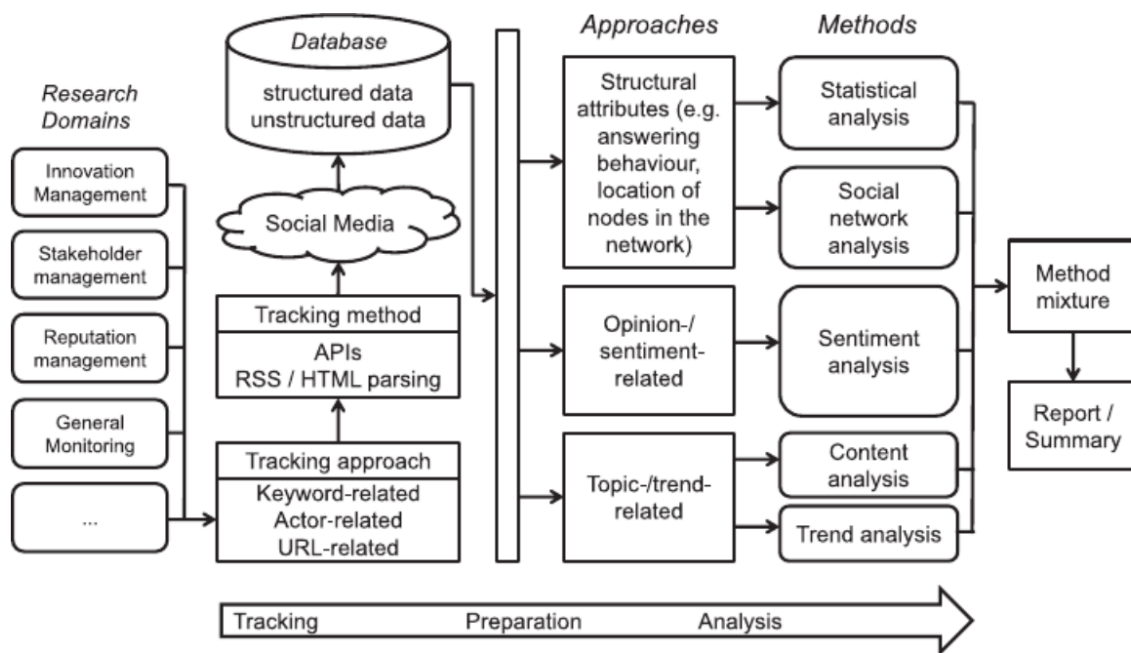
Traditional techniques such as surveys and interviews demand large amounts of time, work and thought. Industries and Governments require solutions to the problem of having to deal with large datasets. Within this unstructured data, there is concealed information which could help the entity understand how exactly the customers or citizens are feeling towards any given subject. Such data can be better managed using the concept of data mining, which is the process of uncovering patterns and significant relationships in large sets of data. Sentiment Analysis is a subset of data mining technology that can be used to extract structured data. This structured data consists of interpretations of the emotions the writer is feeling.

METHODOLOGY

At the first stage, we defined a list of popular topics related to the social tension problem and search queries that are used for obtaining user' posts on these topics. These topics were selected keeping in mind, the current most relevant polarizing social topics in the world. Since, our code consists of a sentiment analyser, and a network model creator; that takes the input of “where on earth identifiers” to access twitter API and the statuses of various IDs; we used different ID's for different topics to maximize the relevancy.

S. Stieglitz et al.

International J



3.1 Collecting posts related to the social tension

Immigration: Since the election of Trump as president, there have been many changes in American foreign relation policy. One of the most contentious being his recent ban on immigration, and how it effects the millions of refugees in the world (with regards to asylum policies) and the American employment and its demographic.

Harassment and Women Empowerment: With the recent surge in the social movement called #metoo, there has been an increase in the awareness of the frequency of physical and emotional assault that people around us, are subjected to on an everyday basis. This is incredibly relevant in India currently, where the movement has recently picked up.

Acceptance of the rights of the LGBTQ: The LGBTQ in a large part of the world, are still considered inhuman and sinners. American recently legalized same-sex marriage in all its 50 states, and India recently recognized the rights of the LGBTQ and legalized

them, we compared the opinions of people on twitter, and off twitter in both India and the world collectively.

Table 1: Search Queries used for the analysis

Topic	Search Query
Immigration and accommodation of refugees	“immigration”
Acceptance of the rights of the LGBTQ	“LGBTQ”
Harassment and Women Empowerment	“#metoo”

3.2 Collecting posts for dataset

We collected tweets, both ‘good’ and ‘bad’ as analysed by our Twitter Analyser to use for further analysis of trends. We gathered around 1000 posts for each of the topics, and then took the top 5 (on the basis of the maximum number of retweets because that corroborates popularity)

Topic	Bad Tweet Example	Good Tweet Example
Immigration and accommodation of refugees	Chinese Refugee Guilty of Violating Immigration Law Gets Probation	The Trump admin. wants families to only apply for asylum at ports of entry—the very places where they're being told to "come...
Acceptance of the rights of the LGBTQ	clearly you are speaking from an extremely selfish point of view since it's common knowledge that there's a direct...	Woooo! @Harry_Styles wins the #GayTimesHonours award for LGBTQ advocate.
Harassment and Women Empowerment	Pamela Anderson slams #MeToo movement and calls feminism boring	if your husband/boyfriend FORCES himself on you it IS #metoo

3.3 Analysis of posts

We analysed the results we got, separately. And then fused the data, to come up with probable future trends and comparisons.

TRACKING

4.1 Keyword Related

Immigration

Table 2: we took out the top 100 tweets on immigration in the US

1	RT @WhiteHouse: The Trump Administration's new rule on immigration is simple, straightforward, and fair: To be eligible for asylum in the U...
2	RT @JRehling: Illegal immigration? "Build a wall!!!" Abortion? "Overturn Roe v. Wade!!!" Gun massacres? "Oh, that's a very complicated pro...
3	RT @zackwhittaker: New: HHS comes clean on its https://t.co/osdMfUfcUQ data breach, confirming that hackers stole income, immigration and t...
4	RT @kwilli1046: Senator Schumer, 2009: Mr Hypocrisy - "Illegal immigration is wrong, plain and simple.... People who enter the United State...
5	RT @tedhesson: The @ACLU filed a lawsuit over the Trump administration's executive actions to block asylum seekers. More on @POLITICOPro: h...
6	RT @FabienDiFilippo: Aide médicale d'Etat : budget encore en augmentation de 53 millions d'€ l'an prochain ! Le Gouvernement ferme les yeu...
7	Again, unconstitutional, violates UN conventions the US has long upheld, deeply racist. Part of worldwide right-win... https://t.co/Yxv0V3wLd8
8	RT @mschlapp: Dems don't believe in our immigration laws and now see that they also don't believe in following election laws. It's called c...
9	RT @sunny_hundal: This is an appalling attack on free speech. Our country allows rich millionaires to claim whatever they want about immig...
10	RT @WyattEarpLA: We saw what they did with Brett. This fight will be dirtier. With RBG on her last leg, they know they must gain control of...

#metoo

Table 3: we took out 100 tweets for #metoo

13	@patcondell Where are the #MeToo / #Femocrats now..... Silent!!!
14	RT @Daflwala: @skhndh @ishaanbhardwaj @hiyer did write his account as part of the #metoo movement. He has shared earlier as well and broug...
15	MeToo விவகாரம்.. Lakshmy Ramakrishnan & Chinmayi வெளியு... https://t.co/4LgTP8iEcX
16	RT @CatherineLEcuye: Una interesante crítica al movimiento #metoo. #yotambien. NPR: Neo-Puritan Revival in the #MeToo Age https://t.co/dbUs...
17	RT @kwhotline: “당신이 바뀔때까지 #METOO 는 멈추지 않는다!” https://t.co/0pD5rBIP
18	While Kubbra Sait backed Nawazuddin Siddiqui, Niharika Singh's #MeToo story mentions Sajid Khan, Bhushan Kumar and... https://t.co/IQzjmPc8DE
19	RT @Diane_Banks: Another Use for AI : Finding Millions of Unregistered Voters 🗳️🗳️ https://t.co/TJglj2ZfCX #resist #notmypresident #trump #res...
20	RT @ksnt63: Fascismo - nazismo razzismo - xenofobia sessismo - maschilismo - omofobia sono un pensiero unico, una sola grandissima montagn...
21	@MicaofColor @ph_phmanuel @Orion0111100110 @Yuma_lein @Ungebeten1 @merdeux_ @Galactica2019 @CABeachpirate... https://t.co/4zqHTw0ZgD
22	RT @caselchris1: What a powerful account. https://t.co/lPjVRHSGtc
23	RT @HavalDarShinde: @bhairavigoswami If your husband/boyfriend FORCES himself on you it IS #MeToo Anyone using violence against you or tou...
24	RT @HavalDarShinde: If your husband/boyfriend FORCES himself on you it IS #MeToo Anyone using violence against you or touching you without...
25	RT @LadyLimps: @LindseyStirling How you can look at the #MeToo and #TimesUp movements & still claim gender equality has been reached If adu...
26	RT @ElizabethLynnCA: Another liberal attack. He calls me racist because I'm white. This guy took it too far with his sexual harassment....
27	While Kubbra Sait backed Nawazuddin Siddiqui, Niharika Singh's #MeToo story mentions Sajid Khan, Bhushan Kumar and... https://t.co/tLVoiGzCd3
28	RT @machine_ml: RT @CyberDeathRay: RT @Diane_Banks: Another Use for AI : Finding Millions of Unregistered Voters 🗳️🗳️ https://t.co/ng7mYbAFY7...

LGBTQ

Table 4: We took out top 100 tweets for LGBTQ in India

6	RT @ClaireHanna: Really lovely to be nominated as Politician of the Year & particularly for LGBTQ Ally of the Year. Well done to all nominee...
7	Proceeds will go toward Jack and Rachel Antonoff's nonprofit the Ally Coalition https://t.co/MJc2ME2cAn
8	What Brazil's election of the far-right's Bolsonaro means for LGBTQ people: https://t.co/lYXZbrPV0i https://t.co/CdR0T9waqH
9	We're taking a look at how #photography is being used to challenge the general understanding of sexuality and gender... https://t.co/hiGFB1XY5C
10	RT @emilylinka: Here's the portrait of @Harry_Styles that's been commissioned for the #GayTimesHonours LGBTQ advocate award. It's on show a...
11	フォロワー100人にご協力ください。あと半分です! 50人ありがとうございます!
12	RT @NotShahem: White gays really are the weak links of the LGBTQ community https://t.co/fqwxKs8HxX
13	I grew up in Harris County. This makes me proud of my hometown & it makes me excited that county youth can see them... https://t.co/lj37tqdFdf
14	RT @wesley_jordan: Just 2 days after the 2018 midterms, Democratic Congressional candidate Justin DeWitt was fired by his employer SJBGroup...
15	Público lgbtq em geral é tudo demente, só uns salvam https://t.co/X99o8WD9Sc
16	RT @LeanneWoodfull: I've teamed up with @BeLong_To LGBTQ+ Youth Services to discuss the importance of speaking up when you're mentally stru...

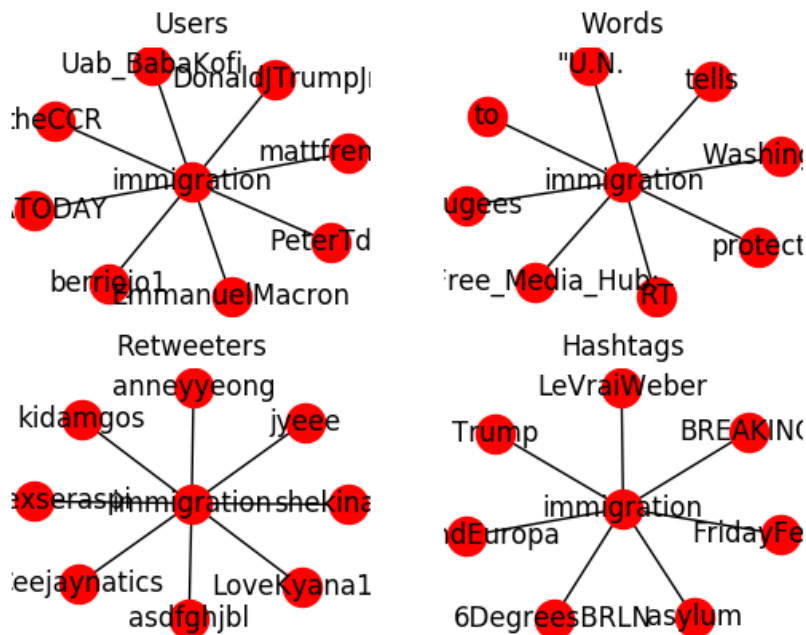
Table 5: we took out 100 top tweets for LGBTQ in the world

	Tweets
1	RT @OfficialWithHL: VIDEO Harry being announced as the winner of the "Gay Times Honour" for LGBTQ Advocate! https://t.co/Ro8oALSG13
2	RT @aintacrow: I wrote this on Black and Pink and @uplcchicago's lawsuit that accuses Illinois DOC of censorship and cutting off LGBTQ pris...
3	RT @RAICESTEXAS: Trump's attacks on asylum directly target trans & other LGBTQ+ people, who aren't safe in parts of Central America. The U...
4	RT @ajplus: Scotland is the first country in the world to make LGBTQ-inclusive school curriculums mandatory. The new policy aims to improve...
5	RT @StevenCanals: PSA: @PoseOnFX is set in the house ball culture, which serves as an alternative family, essential to the survival of blac...
6	RT @GULGBTQplus: We want to tackle discrimination faced by LGBTQ+ people in sport – complete our survey and let us know your experiences an...
7	RT @lukewaltham: It is important that we try to support the LGBTQ+ community in Tanzania in any way possible. Let's sign the petition and s...
8	@xanderberkeley @JamesPMorrison That's why GOP are so often anti-LGBTQ... Self-hate, projection, etc. Look at Mike Pence. Puh-leeze.
9	RT @TexasTribune: Fourteen of the 35 gay, bisexual and transgender candidates who ran for office in Texas during the midterms claimed victo...
10	RT @nypost: Imagine Dragons singer on saving marriage and LGBTQ kids https://t.co/NVMBxKMSGs https://t.co/HVFNSjLEfv
11	More than 200 of the candidates running in congressional and statewide races on Tuesday were black, Latino, Asian A... https://t.co/BVJZBx1Vf3
12	🏳️ Does your world have gay characters? Presumably in the millions of people in it (or even hundreds / thousands... https://t.co/r3MhpIAyFJ
13	RT @RAICESTEXAS: A group of LGBTQ+ migrants in the exodus has banded together under the name La Comunidad. They've faced rampant discrimin...

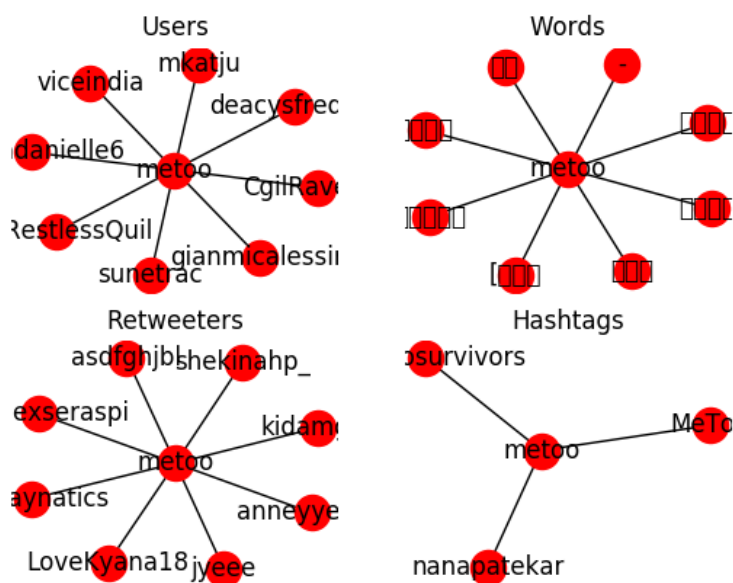
4.2 Actor Related

This section shows the top users, words, retweeters and hashtags related to the aforementioned topics.

Immigration



#metoo



LGBTQ

Table 6: LGBTQ in India

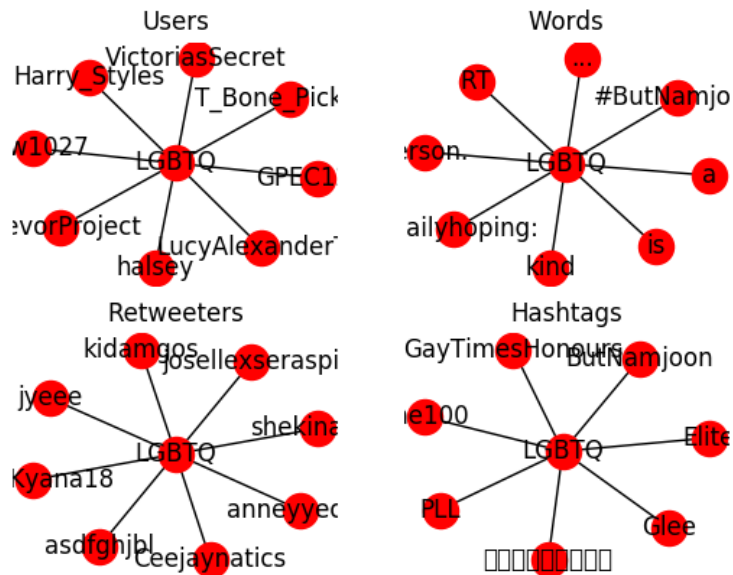
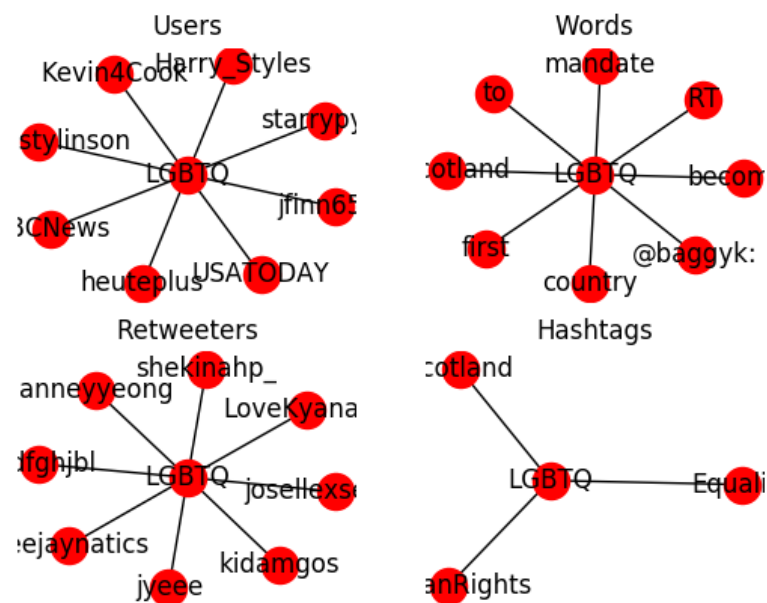


Table 7: LGBTQ in the world



ANALYSIS

6.1 Trend Related

Immigration

- Trump recently suspended some asylum rights, calling illegal immigration a 'crisis' (violates the UN conventions, right wing authoritarianism)
 - Immigration isn't a moral obligation, but a benevolence.
 - Jo Johnson resigns over Theresa May's Brexit plan on trade and immigration, and the EU.
 - Immigration forms are long and hard to fill.
 - ACLU filed a lawsuit against trump's boycott on the asylum situation (President Trump is signing a proclamation declaring that migrants seeking asylum along the southern border must present themselves lawfully at a port of entry.)
 - Anger, that immigration can be dealt with better but isn't being dealt with at all
 - Trump administration is using video calls to speed up the process for migrant children in federal custody
-
- Gave the newest updates on the current news on immigration
 - Not only immigration related news on the US
 - Also gave some unbiased facts about the immigration process

(gives me an idea that twitter has all 3 different kinds of news)

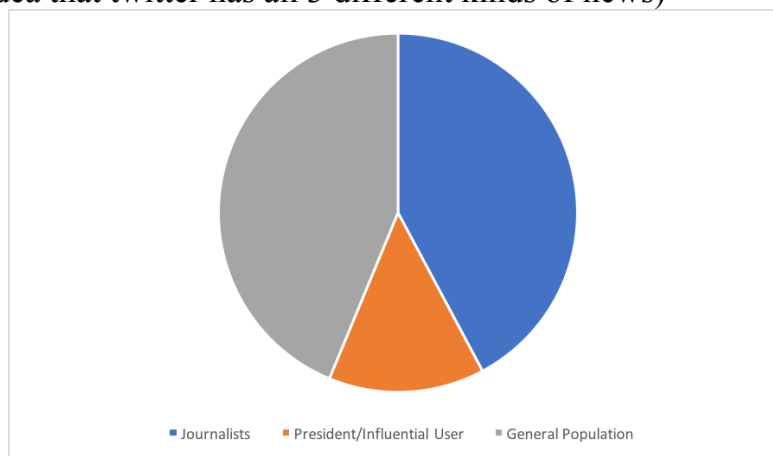


Table 8: depiction of user demographic for network graphs for 'immigration'

#metoo

From the list of 30 tweets one can clearly see that India is quite on edge when it comes to allegations against Bollywood actors like Nana Patekar or Nawazuddin Siddiqui and many of these tweets are from their co actors like Salman Khan and Suddhan Sadaf

rising against the allegations and simply calling out the allegations and calling them baseless.

Suknidh Kaur, a very renowned poet, is also conducting a metoo meet in Chennai with her followers in order to talk about and address it with the audience. A few tweets mentioned how metoo is not quite prevalent in a country like India where acts like marital rapes aren't an actual offence in India and there were concerns about the hypocrisy of the same. These tweets created a huge buzz with the tweeters.

Feminists were targeted in a particular case wherein #HimToo was being made into a movement and men came out to talk about their experiences with sexual assault and similar cases where they were the targets. A case where a twitter user questioned if parents could blame their child for an assault. This created quite an uprising and was retweeted several times. Some people were raising awareness about the fact that possibly a large chunk of the society today is oblivious to the real concept of #MeToo and would rather have them understand the true urgency of it.

Although there were a couple of tweets from people belonging to China and Spain who appreciated how the #MeToo movement had reached India as well, there was one by Elizabeth Lynn where she called out an Indian man she met on her trip here who allegedly called her a racist for the simple reason that she was a foreigner and had a fair complexion. While all these raised concerns among the Twitter community there were a tweet which trivialised the entire movement by making a joke out of the concept by indulging in a play with words.

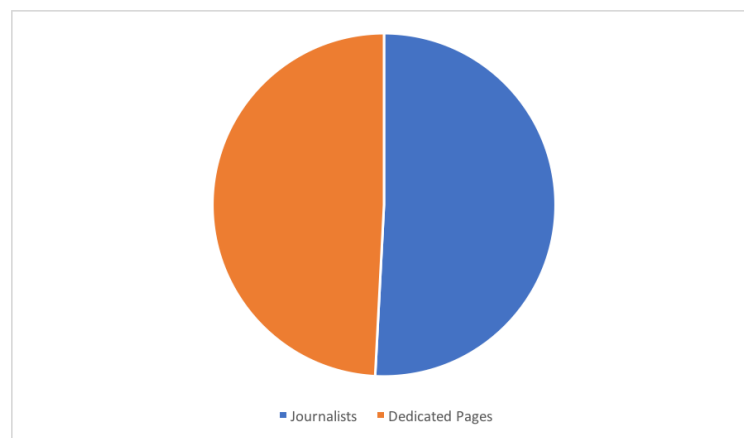


Table 9: depiction of user demographic for network models in '#metoo'

LGBTQ

India:

- Blacks feel like the white LGBTQ are the weakest link of this community
- 17 black women elected as judges including 1 LGBTQ woman in a Texas county
- Links of photography challenging how gender is perceived

- Scotland is the first country in the world to make the LGBTQ inclusive school curriculum mandatory
- Harry Styles winning the Gay Times Honor award for being an LGBTW Advocate
- Lady Gaga news about her standing up for the LGBTQ and creating awareness
- Shared pictures of trans models
- LGBTQ candidates in political positions all over the world

World:

The first 30 tweets talk about celebrities coming into the limelight as gay pop icons, in most cases, Harry Styles, who was announced as the winner by the Gay Time Honour. A few tweets appreciated that Scotland had become the first country to have LGBTQ+ inclusive schools in the world.

Trump became the centre of controversies again when his attacks on asylum directly targeted the trans and the LGBTQ+ community. There were many petitions being circulated to support the rights of the LGBTQ+ community in several countries like Tanzania and to stop any kind of discrimination against them.

There was a few highlighting how 35 gay, transgender and bisexual candidates claimed victory in Texas in the midterm elections which took place in the USA and more than 150 people belonging to the LGBTQ community were voted into office on Tuesday. Many claimed that a rainbow tsunami had flooded the world. There were a couple of haters who voiced their opinion on this “unnatural” way of being and calling the world foolish to be progressing in this direction. They said the world was collectively headed towards doom.

6.2 Opinion and Sentiment Related

Immigration

- Good tweets percentage: 33 %
- Bad tweets percentage: 23 %
- Unbiased tweets percentage: 42 %

Good Tweets (there is an assumption that immigration is good, and those people should have rights)

- ACLU to challenge Trump’s new asylum ban
- Trump’s new rule on immigration
- Rich millionaires claiming they know everything about immigration, and claiming and being able to make the rules
- All about the recent proclamation of the asylum rules for refugees

- People shouldn't conflate gang violence like the MS13 with immigration, because they aren't the same thing. Those politicians are culprits of the same problem.

Basic Idea:

People should fight for the rights of the refugees, and there should be better awareness of what it means to be one, and their plight.

There is also a call to focus on the internal issues in the US, and not just the bigger more interesting ones.

Bad tweets

- Trump's new asylum ban
- The new immigration decree is a farce
- The unknown future for the migrant children
- The suing of the trump ban by ACLU

Basic Idea:

It talks about how unknown the future for the migrant children now in the US is, and that there has been largely a resistance with regards the Trump's asylum ban

#metoo

- Good tweets percentage: 27 %
- Bad tweets percentage: 8 %
- Unbiased tweets percentage: 64 %

Good Tweets

It's a comparison how India is better off than a few countries like Japan where sexual harassment isn't considered a crime. There are many users who have appreciated strong women who have managed to muster up enough courage and stand up against their offenders.

Bad Tweets

These are mostly people giving a narrative of their instances and experiences with sexual harassments and assaults. A few of them voice their disappointment towards the news of Nawazuddin Siddiqui being involved on the basis of him being an extremely versatile actor. Pamela Anderson trivialised the concept by calling it boring and completely slamming the third wave of the feminism movement.

LGBTQ

India:

- Good tweets percentage: 35 %

- Bad tweets percentage: 15 %
- Unbiased tweets percentage: 50 %

Good Tweets:

- Lady Gaga speaking up on health, sexual assault, and the LGBTQ
- Scotland is the first country in the world to make the LGBTQ inclusive school curriculum mandatory
- Explain to the people who don't understand, what LGBTQ is
- Brazil's election with a fair right president, and what it means for the LGBTQ
- Claire Hanna (MLA for South Belfast) was nominated as politician of the year, but also LGBTQ Ally of the year

Basic Idea:

Showed the most recent achievements and hurdles the LGBTQ movement has overcome. Talks about the effects of recent laws being passed, to the community.

Bad Tweets:

- Injustice done to the LGBTQ youth growing up in the orthodox faith
- Black men of the LGBTQ community, calling out white men on being the weak links
- Victoria's Secret models not taking people from the LGBTQ as their audience is mostly white men who won't appreciate it
- News about parents feeling like they've failed, if their child is LGBTQ

Basic Idea:

Addresses the unjust to various demographics within the community, and some internal negative discourse.

World:

- Good tweets percentage: 59 %
- Bad tweets percentage: 6 %
- Unbiased tweets percentage: 33%

Good Tweets:

Many of the twitter users were elated about Harry Styles being awarded the Gay Times Honour and thought that he deserved it, while another singer/celebrity was appraised for using her platform to spread awareness about issues like mental health and the LGBT community. Scotland has come to the forefront with making schools curriculums more inclusive in order to balance the rights of the LGBTQ+ community as well. Many

people were extremely glad that about 150 people of the LGBTQ emerged victorious in the midterm elections in the USA.

Bad Tweets:

There are many places in the world where same sex marriages and being in love with the same gender are still considered to be against the law. One such country is Zanzibar where 10 men were arrested for conducting a same sex marriage. People look down upon the LGBTQ and empathise with those who are pro to this obscene concept and go ahead to say that humanity is up for doom if we continue to tread this path.

6.3 Structural Attributes

Immigration

Small Word for Tweets

Degree Centrality	1- Emmanuel Macaron (French President) 4- Daniel Collier	Both have a lot of ties, and are important in the first stage of message transferring (has a lot of friends)
Closeness Centrality	1- Emmanuel Macaron	Best and fastest way to interact with a lot of masses (people retweet hi stuff)
Betweenness Centrality	1- Emmanuel Macaron	Has power, in controlling the flow of information and serves as an important link between Reuters, USATODAY (new agencies) and the people

Random World for tweets

We see 3 different communities, that aren't inter connected.

This shows the degree of opinion bubbles that can form on social media.

0-4

USATODAY is an American news agency, and Collier is an American citizen that subscribes to it. He's getting all his immigration input from USA Today, which means he's going to believe whatever they say. (USATODAY is right biased)

6-5

Reuters, DStreak1

DStreak1 is another American citizen, who gets his immigration news from Reuters. We can assume he will not be biased as Reuters is an unbiased news agency.

1,2,3,7

This set includes the French president, who clearly has a high centrality, 2 normal American citizens and the fox news channel.

This tells me that fox news, will largely be right biased, as it aligns with Macron's views. This shows heavy influence of the views of the French president in the news agency, and the opinion bubble that hence gets formed in the minds of the tweeters.

Degree Centrality	1- Emmanuel Macron	Has a lot of ties, and are important in the first stage of message transferring
Closeness Centrality	1- Emmanuel Macron	Best and fastest way to interact with a lot of masses (people retweet hi stuff)
Betweenness Centrality	1- Emmanuel Macron	Has power, in controlling the flow of information and serves as in important link

Small world for Retweeters

All of the nodes seem to be connected in a news agency-> reader fashion, which indicates that all users mostly retweet the stuff of news channels. The retweets of the news channels, is largely from large personalities like the French resident.

Degree Centrality	6- Dstreak1	Has a lot of ties, and are important in the first stage of message transferring
Closeness Centrality	5- Reuters 6-Dstreak1	Best and fastest way to interact with a lot of masses (people retweet his stuff)
Betweenness Centrality	6- Reuters	Has power, in controlling the flow of information and serves as in important link

Random world for Retweeters

(0,7,3)-1

This shows a heavy retweeting by the president from these news agencies, which are mostly rightist. This will have heavy impact on the people that follow Macron, which are a lot since his centrality id definitely the highest.

User 2,6 seem to be the mediators between rightist ideology coming from macron, and the more neutral content coming from Reuters and some other users.

Degree Centrality	1- Emmanuel Macaron	Has a lot of ties, and are important in the first stage of message transferring
Closeness Centrality	1- Emmanuel Macaron	Best and fastest way to interact with a lot of masses (people retweet his stuff)
Betweenness Centrality	1 –Emmanuel Macaron	Has power, in controlling the flow of information and serves as in important link

#metoo

Small world for Tweets

Degree Centrality	3- karinadanielle6 4- TheRestlessQuil 6- France3tv	These are journalists, authors and activists based in France, every one contributing to the AsiaBibi case and how Indians are dealing with empowered women (depicts connections).
Closeness Centrality	3- karinadanielle6 4- TheRestlessQuil	Being a journalist and activists respectively a lot of people are inspired from them and retweet quite a lot of their opinions on #MeToo and #AsiaBibi
Betweenness Centrality	3- karinadanielle6	Being a journalist based in France, she holds immense power, in controlling the flow of information and serves as in important link between the movement and linking it to cases like #AsiaBibi

Random World for tweets

There are 2 broad communities here, according to the network graph.

3 – karinadanielle6

A writer based in France who has worked quite actively on the #AsiaBibi case and helped her and her family stay protected and also to raise her voice against this exhaustive imprisonment for being a Christian.

1,2,4,5,6,7

Ananya Mukherjee, Gian Micalessin, Markanday Katju are all activists based in India and are constantly waging a war against the people who target people and sexually harass them. France3tv, PublicSexCrimes, TheRestlessQuil and viceindia are all journals or social media platforms dedicated to the #MeToo movement and helping people in rising above their fears and raise their voices against assaults and harassments.

Degree Centrality	0- viceindia	Has collected a lot of information from all sources in India, and are important in the first stage of passing the information.
Closeness Centrality	0- viceindia	Best and fastest way to interact with most people in India.
Betweenness Centrality	0-viceindia	They hold the power and monopoly in controlling the flow of information in India. Most of the foreign journalists also take their information from here.

Small World for Retweeters

Degree Centrality	2- gianmicalessin	Italian journalist and activist working for women rights in India, and keeps a track of the changes in the country.
Closeness Centrality	5- PublicSexCrimes	A page dedicated for the movement and is the best and fastest way to interact with most people.
Betweenness Centrality	5- PublicSexCrimes	They hold the power and monopoly in controlling the flow of information. Most of the people come to this page to get updates on the #MeToo movement.

Random world for Retweeters

The community of retweeters is categorised into 2 parts. 0-6 and then 7.

Ananya Mukherjee is a journalist and nationalist in India and she has raised her voice several times against such assaults.

Degree Centrality	1 – mkatju	He has a lot of ties being a controversial judge in India and the chairman of the Press Council of India, and is important in the first stage of message transferring
Closeness Centrality	1-mkatju	Best and fastest way to interact with a lot of people being the chairman of the Press Council of India, he is immensely active with the on goings in India
Betweenness Centrality	1-mkatju	He clearly holds a lot of power, in controlling the flow of information and serves as in important link

LGBTQ

India:

Small World Network for Users

(0-5)

5 is a general user, who is a harry styles fan and follows him. So, as a result, she gets the little knowledge she coincidentally gathered on the LGBTQ community.

(3,6,7)-1

7 has a betweenness centrality of 0 and is a part of this graph because, 5,6 are both following BNC news and are gamers.

Degree Centrality	0- Harry Styles 1- NBC News	Both have a lot of ties, and are important in the first stage of message transferring (has a lot of friends)
Closeness Centrality	1- NBC News	Best and fastest way to interact with a lot of masses (people retweet their stuff)
Betweenness Centrality	1- NBC News	Has power, in controlling the flow of information and serves as in important link between harry styles fans, and

gamers (which are very different people)

Random Network for Users

7-6

6 is a random citizen, who has access to information on the LGBTQ through the radio channel.

The rest of the network is very densely connected, and IGDA is a gaming association that is linked as 5 enjoys both harry styles and pop culture, and also gaming.

Degree Centrality	4- New1027 7- IGDA	Both have a lot of ties, and are important in the first stage of message transferring (has a lot of friends)
Closeness Centrality	4-New1027 7-IGDA	Best and fastest way to interact with a lot of masses (people retweet their stuff)
Betweenness Centrality	7- IGDA	Has power, in controlling the flow of information and serves as in important link

Small World Network for Retweeters

Degree Centrality	2-MortgageMogul	Has a lot of ties, and are important in the first stage of message transferring (has a lot of friends)
Closeness Centrality	2- MortgageMogul	Best and fastest way to interact with a lot of masses (people retweet their stuff)
Betweenness Centrality	2- MortgageMogul	Has power, in controlling the flow of information and serves as in important link

Random Network for Retweeters

This graph shows that there are some major opinion bubbles, that don't interact with another.

Degree Centrality	0- Harry Styles	Harry Styles interacts with NBC, and retweets from NBC and Mortgage Mogul takes from Harry Styles
Closeness Centrality	0- Harry Styles	Best and fastest way to interact with a lot of masses (people retweet their stuff)
Betweenness Centrality	0- Harry Styles	Has power, in controlling the flow of information and serves as in important link

World:

Small world for Tweets

Degree Centrality	2- angelhrndzz 7- NuevoEnNetflix	These are extremely relevant and woke youtubers and pages attempting to change the world in putting out content in more pleasing manner (depicts connections).
Closeness Centrality	2 – angelhrndzz 7 – NuevoEnNetflix	Extremely vocal in their beliefs and putting it out in a more pleasing manner, people are inspired by them and retweet quite a lot of their opinions on the LGBT.
Betweenness Centrality	2 – angelhrndzz 7 – NuevoEnNetflix	They are linked to the LGBTQ+ in a more uninhibited manner and play a pivotal role in controlling the flow of information and serves as in important link.

Random World for Tweets

There are 3 communities in the network – (0,1,2,6,4) (5) and (3,7)

The largest community is of the popular woke celebrities like Harry Styles and Angel Hrndzz, whom people follow and derive their inspiration and source of information on the LGBTQ community. While Harry Styles and Angel Hrndzz are more of gay icons, the others like NYPost, NBCNews and PandlerNews are news channels which provide accurate and relevant information and updates.

The second community comprises of producers and thinkers who put out the idea of a different inclusive society by means of entertainment like media, youtube, movies and theatre. These are extremely relevant in today's fast paced world where not every individual watches/reads the news and prefers to derive their information from such sources.

The third network is a Harry Style appreciation page by people belonging to the LGBTQ+ community and a couple of fangirling teenagers who have immense love for the greatest pop gay icon today.

Degree Centrality	2 – angelhrndzz	Has collected a lot of information from all sources on LGBTQ and is important in the first stage of passing the information.
Closeness Centrality	2 – angelhrndzz	Best and fastest way to interact with most people over the world.
Betweenness Centrality	2 – angelhrndzz	Holds major monopoly in the passage of information between the journalists and the unaffected society and the gay community.

Small World for Retweeters

Degree Centrality	3- WynonnaEarP 4- OwreyKris	These are TV series and an artist which have derived inspiration from the LGBTQ community and hence keep a track of the changes in the world.
Closeness Centrality	4- OwreyKris 6-AbacusMike	A page by artists/comic artists who are dedicated for the movement and is the best, fastest and easiest way to interact with most people.
Betweenness Centrality	4- OwreyKris 6-AbacusMike	They hold the convenience and ease in controlling the flow of information. Most of the people come to this page to get updates on the movement.

Random Network for Retweeters

It is a highly community in itself, and every node is connected in some way or the other and there exists only a single community of retweeters.

Degree Centrality	5 – onlyayangel	Being an immense Harry Styles fan and also belonging to the LGBTQ community, she has a lot of ties and is important in the first stage of message transferring
Closeness Centrality	5 – onlyayangel	She has a lot of people retweeting her and she is quite woke on the LGBTQ+ in general and hence is one of the best and fastest way to interact with a lot of people collectively.
Betweenness Centrality	5 – onlyayangel	She clearly holds a lot of power, in controlling the flow of information and serves as an important link between the gay community and the other audience.

RESULT AND CONCLUSION

Immigration:

- There seems to be a general shift in the right-wing belief within Americans, to not want immigrants and refugees into the United States.
- Most international bodies are fighting against the Trump Administration's policies, and are against them.
- The French president, is highly active on twitter, and is a strong link between various News Channels and the general public.
- There were a lot of opinion bubbles, that were forming amongst the conservative and liberal views, which ran parallel to the real-world scenario.

#metoo

- There is a lot of undue aversion towards the women who have attempted to stand up against their offenders, who also happen to belong to the Bollywood industry.
- There are many popular celebrities trying to gather more support and make the people aware of the campaign and what it stands for.
- There was discord from the males of the society demanding their due of 'victimisation'.
- Considering that marital rape is not even a criminal offence in a country like India, the entire basis of the #MeToo movement was slashed and trivialised.
-

LGBTQ

7.2 Top 30 tweets

While, both in India and around the world the common topics seemed to be the recent inclusions of the LGBTQ in school curriculums, there were many differences:

India	World Wide
<ul style="list-style-type: none">• Black Women being elected as judges• Lady Gaga• Photography• Black men calling out white LGBTQ for being the weak links	<ul style="list-style-type: none">• Trump (attack of asylum)• Tanzania raising voices for the Gay community• 35 LGTQ- 35 people won in mid term• LGGBTQ is unnatural, and we're wasting time- we're collectively heading towards doom (-)

Analysis:

- The big developments, were prevalent irrespective of country
- World was talking more on trump, which India wasn't too concerned about

7.3 Sentiment Analysis

India	World
Good tweets percentage: 35 %	Good tweets percentage: 46 %
Bad tweets percentage: 15 %	Bad tweets percentage: 12 %
Unbiased tweets percentage: 50 %	Unbiased tweets percentage: 40 %

Analysis:

- India had more bad tweets on the LGBTQ
- Also, they had less good
- More unbiased
- When something is more unbiased, but there less good tweets, goes to show that people in India have more hard-line and orthodox stances on the LGBTQ as compared to the rest of the world.

Good Tweets

- While worldwide, the conversation seems more focused on the achievements of the people of the LGBTQ community, or its advocates, the ones specific to India were discussing internal mechanisms of changes taking place with regards to policies that will affect this minority community. e.g. Jair, in Brazil, and what it generally means to be LGBTQ in today's world.
- We accrue this to the lessening of specifications as we increase our survey pool.
- But also, it's fair to say given previous deductions that India is still in more internal conversation about who the LGBTQ as compared to the rest of the world.

Bad Tweets

- While in the world, there was calling out of the LGBTQ community being spineless and weak, and also not knowing how to identify as this community but just doing it because it was 'cool'. The world questions, the authenticity of the belief of these people.
- In India, the conversation was on similar lines where people felt that they've failed as parents if their child wasn't straight. There was a lot of conversation of Victoria's secret turning down trans models, and how the trans youth would grow up with parents of an orthodox faith.
- There were parallels, and the concerns are similar; we see the major question being of authenticity of belief.

7.4 Top 5 tweets

Analysis:

- In the world, the top tweets were US centric and revolving around Trump and the mid-terms.
- India, seemed to be talking more internal and sharing various specific incidences of the way that the lives of the LGBTQ are governed on an everyday basis.

7.5 Network Graphs

Small word tweets

India	World
NBC News	blogger, who belongs to the LGBTQ community, doesn't censor, puts content in a more pleasing manner

Analysis:

World listens to a blogger, India listens to the news

Random World tweets

India	World
IDGA (gaming association)	(More accurate) blogger, who belongs to the LGBTQ community, doesn't censor, puts content in a more pleasing manner

Analysis:

- The world is clearly divided into 3 communities namely, the woke, the entertainers, and fan pages.
- This shows the formation of opinion bubbles, where the users don't interact with people of different opinions.
- India, was more densely connected and showed variation in the likes of people, this meant the gamers were also harry styles fans. This allowed a better propagation of opinion, and less opinionated bias.

Small world retweets

India	World
A Mom	Comic Artist

Analysis: Similar, people take interest in the current social movements.

Random world retweets

India	World
Harry Styles	Harry Styles fan- reposting a lot of Harry Styles so took interest because of him

Analysis: people in India don't retweet as much.

Conclusion:

- There is still, hesitation in India with respect to LGBTQ
- More internal mechanism discussion, as there isn't immense clarity on what the term means and who these people are.
- More emphasis on censorship in India, as the Twitter uses mostly repost from the news over bloggers and activists etc.
- General overall acceptance (policies) as there is support for any progressive laws passed in the favour of the LGBTQ community.
- World blames the person, India blames the parent.

7.6 Comparative Analysis of the Twitter community in India, and the community around us

Do you have a Twitter account?

107 responses

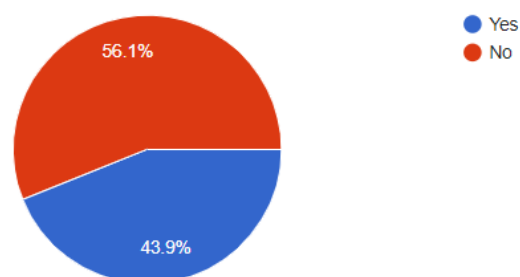
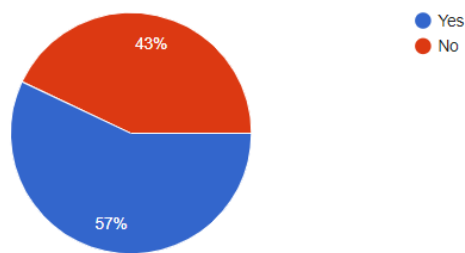


Table 10: This question is just to depict our demographic, wherein 43.9% of our survey pool was on Twitter.

Do you believe you fully understand the plight of the LGBTQ?

107 responses

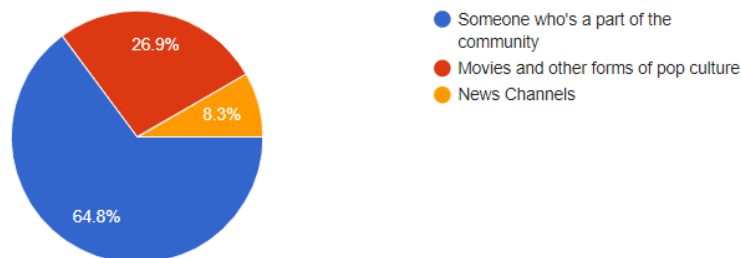


Quite contrary to the analysis obtained from the twitter content, most of the people believed that they understood the plight of the people belonging to the LGBTQ community and did not actually require the need for the intense internal discussion to formulate their basic understanding of the same.

From the twitter analysis, however, it was established that most Indians were still indulging in a lot of internal discussion over understanding their basic concerns and rights.

Your preferred channel to learn all things LGBTQ?

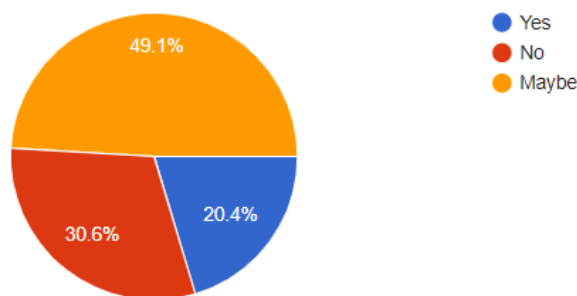
108 responses



Unlike the data obtained from the twitter analysis, where most people were sourcing their information off the biased news channels, most people surveyed said that they preferred to hear about it and then form their understanding about the same from someone who is actually a part of the LGBTQ community. Following which, there is a considerable chunk of our demographic which watches and learns about the community from more unbiased means like the internet, blogs, theatre and the movies. But the smallest majority actually receives their information from news channels.

Would your parents be okay with you not being straight?

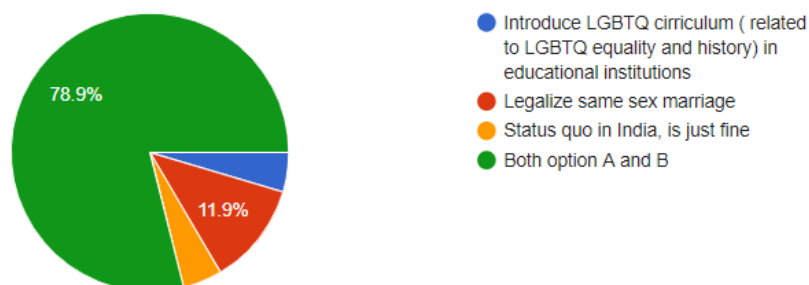
108 responses



With the popular belief among the twitter users that most parents are to blame for raising their child wrong if they have queer preferences, and when people were asked if they thought if their parents if they will be progressive enough to accept them as they are, most of the survey population was uncertain about the same and about half of them replied with a maybe. The smallest percentage replied with a yes, which goes ahead to show that most people are still hesitant with their identities.

With respect to the LGBTQ community:

109 responses



Although, this response is probably quite short sighted considering that they wish to be making such conclusively large changes for people whose acceptance in the society is still quite debatable. However, this shows that most of the youth today (most of our demographic) are accepting and accommodating of people's individual choices and beliefs.

REFERENCES

- [1] The rate of reply and nature of responses to suicide-related posts on Twitter
(<https://www.sciencedirect.com/science/article/pii/S2214782917301008>)
- [2] Identifying Fake News and Fake Users on Twitter
(<https://www.sciencedirect.com/science/article/pii/S1877050918312559>)
- [3] A picture tells a thousand words: What Facebook and Twitter images convey about our personality (<https://www.sciencedirect.com/science/article/pii/S0191886916312478>)
- [4] Human Mobility Prediction Through Twitter
(<https://www.sciencedirect.com/science/article/pii/S1877050918311165>)
- [5] Using Tweets Embeddings For Hashtag Recommendation in Twitter
(<https://www.sciencedirect.com/science/article/pii/S1877050918301030>)
- [6] An Ensemble Classification System for Twitter Sentiment Analysis
(<https://www.sciencedirect.com/science/article/pii/S187705091830841X>)
- [7] Detection of spam-posting accounts on Twitter
(<https://www.sciencedirect.com/science/article/pii/S0925231218308798>)
- [8] Analysis of Comments of Users of Social Networks to Assess the Level of Social Tension
Level of Social Tension by *Dmitry Donchenko, Nadezhda Ovchar, Natalia Sadovnikova, Danila Parygin, Olga Dmitry Donchenko, Nadezhda Ovchara, Natalia Sadovnbikova, Danila Parygin*
- [9] Social Network and Sentiment Analysis on Twitter: Towards a Combined Approach by
Paolo Fornacciari, Monica Mordonini, Michele Tomauiolo from Dipartimento di Ingegneria
dell'Informazione Università degli Studi di Parma Parma, Italy

APPENDICES

9.1 Code used for implementation

```
import twitter #access twitter
import pprint
import json #javascript object notation
from pprint import pprint
import os
import re
import tweepy #to access twitter API
from tweepy import OAuthHandler
from textblob import TextBlob #to measure polarity and subjectivity
from googleapiclient.discovery import build
from collections import Counter
from prettytable import PrettyTable #creating tables
import matplotlib
matplotlib.use('TkAgg')
import matplotlib.pyplot as plt
import easygui #gui for the project
import pprint
import sys
import networkx as nx
import matplotlib.pyplot as plt
import pandas as pd
from IPython.display import display
from texttable import Texttable

WORLD_WOE_ID = 23424848
USA_WOE_ID= 23424977
INDIA_WOE_ID= 23424848
def twitter_setup():
    consumer_key = '2D2qQNR2rSUwzulwEGFPXR3zp'
    consumer_secret = 's7A2pwgTvo0dZvzbA5lzP06YG9tm3IST9qZjkAkF2spKrXxyW1'
    access_token = '1007573251231383553-nu10kcCDaJDLd9UDVsVZx6FFKy5BKo'
    access_token_secret = '2L10gEakQnGeZe4g88d13mgywW6e0oRYkiBN8kFhvixE1'
    auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
    auth.set_access_token(access_token, access_token_secret)
    api = tweepy.API(auth)
    return api

class TwitterClient(object):
    def __init__(self):
        consumer_key = '2D2qQNR2rSUwzulwEGFPXR3zp'
        consumer_secret = 's7A2pwgTvo0dZvzbA5lzP06YG9tm3IST9qZjkAkF2spKrXxyW1'
        access_token = '1007573251231383553-nu10kcCDaJDLd9UDVsVZx6FFKy5BKo'
        access_token_secret = '2L10gEakQnGeZe4g88d13mgywW6e0oRYkiBN8kFhvixE1'
        try:
            self.auth = OAuthHandler(consumer_key, consumer_secret)
            self.auth.set_access_token(access_token, access_token_secret)
            self.api = tweepy.API(self.auth)
        except:
            print("Error: Authentication Failed")

    def clean_tweet(self, tweet):
        return ' '.join(re.sub("(@[A-Za-z0-9]+)|([^0-9A-Za-z \t])|(\w+:\/\/\S+)", " ",
tweet).split())

    def get_tweet_sentiment(self, tweet):
        analysis = TextBlob(self.clean_tweet(tweet))
        if analysis.sentiment.polarity > 0:
            return 'positive'
        elif analysis.sentiment.polarity == 0:
            return 'neutral'
        else:
            return 'negative'
```

```

def get_tweets(self, query, count=10):
    tweets = []

    try:
        fetched_tweets = self.api.search(q=query, count=count)
        for tweet in fetched_tweets:
            parsed_tweet = {}
            parsed_tweet['text'] = tweet.text
            parsed_tweet['sentiment'] = self.get_tweet_sentiment(tweet.text)
            if tweet.retweet_count > 0:
                if parsed_tweet not in tweets:
                    tweets.append(parsed_tweet)
            else:
                tweets.append(parsed_tweet)
        return tweets

    except tweepy.TweepError as e:
        print("Error : " + str(e))

def view_trends():
    CONSUMER_KEY = '2D2qQNR2rSUwzulwEGFPXR3zp'
    CONSUMER_SECRET = 's7A2pwgTvo0dZvzbA5lzP06YG9tm3IST9qZjkAkF2spKrXxyW1'
    OAUTH_TOKEN = '1007573251231383553-nu10kcCDaJDLd9UDVsVZx6FFKy5BKo'
    OAUTH_TOKEN_SECRET = '2L10gEAKQnGeZe4g88d13mgYW6e0oRYkiBN8kFhvixE1'
    auth = twitter.oauth.OAuth(OAUTH_TOKEN, OAUTH_TOKEN_SECRET, CONSUMER_KEY,
    CONSUMER_SECRET)
    twitter_api = twitter.Twitter(auth=auth)
    world_trends = twitter_api.trends.place(_id=WORLD_WOE_ID)
    print("-" * 50)
    print('World Trends : ')
    print("-" * 50)
    for i in world_trends[0][u'trends']:
        print(i[u'name'])
    sys.stdin.read(1)

def change_trends():
    global WORLD_WOE_ID
    data_enterbox_ID=easygui.enterbox('Enter the WOEID to change location: ', 'WOEID
LOCATION WISE')
    WORLD_WOE_ID = int(data_enterbox_ID)
    CONSUMER_KEY = '2D2qQNR2rSUwzulwEGFPXR3zp'
    CONSUMER_SECRET = 's7A2pwgTvo0dZvzbA5lzP06YG9tm3IST9qZjkAkF2spKrXxyW1'
    OAUTH_TOKEN = '1007573251231383553-nu10kcCDaJDLd9UDVsVZx6FFKy5BKo'
    OAUTH_TOKEN_SECRET = '2L10gEAKQnGeZe4g88d13mgYW6e0oRYkiBN8kFhvixE1'
    auth = twitter.oauth.OAuth(OAUTH_TOKEN, OAUTH_TOKEN_SECRET, CONSUMER_KEY,
    CONSUMER_SECRET)
    twitter_api = twitter.Twitter(auth=auth)
    world_trends = twitter_api.trends.place(_id=WORLD_WOE_ID)

    print('\n Altered Trends : \n')
    for i in world_trends[0][u'trends']:
        print(i[u'name'])
    sys.stdin.read(1)

def combined_graph():
    CONSUMER_KEY = '2D2qQNR2rSUwzulwEGFPXR3zp'
    CONSUMER_SECRET = 's7A2pwgTvo0dZvzbA5lzP06YG9tm3IST9qZjkAkF2spKrXxyW1'
    OAUTH_TOKEN = '1007573251231383553-nu10kcCDaJDLd9UDVsVZx6FFKy5BKo'
    OAUTH_TOKEN_SECRET = '2L10gEAKQnGeZe4g88d13mgYW6e0oRYkiBN8kFhvixE1'
    auth = twitter.oauth.OAuth(OAUTH_TOKEN, OAUTH_TOKEN_SECRET, CONSUMER_KEY,
    CONSUMER_SECRET)
    twitter_api = twitter.Twitter(auth=auth)
    q = input('Enter a Name to Search: ')
    print('Extracting Data');
    search_results = twitter_api.search.tweets(q=q, count=70)
    statuses = search_results['statuses']
    hashtags = [hashtag['text']
        for status in statuses
        for hashtag in status['entities']['hashtags']]

```

```

retweets1 = twitter_api.statuses.retweets(id=317127304981667841)
status_texts = [status['text']
                 for status in statuses]
words = [w
          for t in status_texts
          for w in t.split()]
print('Processing Data');
for _ in range(5):
    try:
        next_results = search_results['search_metadata']['next_results']
    except KeyError, e:
        break
    kwargs = dict([kv.split('=') for kv in next_results[1:].split("&")])
    search_results = twitter_api.search.tweets(**kwargs)
    statuses += search_results['statuses']
screen_names = [user_mention['screen_name']
                 for status in statuses
                 for user_mention in status['entities']['user_mentions']]
screen_names=set(screen_names)
screen_names=list(screen_names)
G=nx.Graph()
for r in min(screen_names,screen_names[0:8]):
    G.add_edge(str(q), str(r));
plt.subplot(2,2,1);
plt.title('Users')
nx.draw(G, with_labels = True)
#plt.savefig('social/'+q+'1.png')
#plt.show(block=True)
#plt.close('all')
G2=nx.Graph()
for r in min(words,words[0:8]):
    G2.add_edge(str(q),r);
plt.subplot(2,2,2);
plt.title('Words')
nx.draw(G2, with_labels = True)
#plt.savefig('social/'+q+'2.png')
#plt.show(block=True)
#plt.close('all')
G3=nx.Graph()
for r in min(retweets1,retweets1[0:8]):
    G3.add_edge(str(q), str(r['user']['screen_name']));
plt.subplot(2,2,3);
plt.title('Retweeters')
nx.draw(G3, with_labels = True)
#plt.savefig('social/'+q+'3.png')
#plt.show(block=True)
#plt.close('all')
G4=nx.Graph()
for r in min(hashtags,hashtags[0:8]):
    G4.add_edge(str(q), r);
plt.subplot(2,2,4);
plt.title('Hashtags')
nx.draw(G4, with_labels = True)
plt.savefig('social/'+q+'4.png')
plt.show(block=True)
plt.close('all')

def network_models():
    CONSUMER_KEY = '2D2qQNR2rSUwzulwEGFPXR3zp'
    CONSUMER_SECRET = 's7A2pwgTvo0dZvzbA5lzP06YG9tm3IST9qZjkAkF2spKrXxyW1'
    OAUTH_TOKEN = '1007573251231383553-nu10kcCDaJDLd9UDVsVZx6FFKy5BKo'
    OAUTH_TOKEN_SECRET = '2L10gEAKqnGeZe4g88d13mgywW6e0oRYkiBN8kFhvixE1'
    auth = twitter.oauth.OAuth(OAUTH_TOKEN, OAUTH_TOKEN_SECRET, CONSUMER_KEY,
CONSUMER_SECRET)
    twitter_api = twitter.Twitter(auth=auth)
    q = input('Enter a Name to Search: ')
    print('Extracting Data');
    search_results = twitter_api.search.tweets(q=q, count=30)
    statuses = search_results['statuses']
    retweets1 = twitter_api.statuses.retweets(id=317127304981667841)

```

```

print('Processing Data');
for _ in range(5):
    try:
        next_results = search_results['search_metadata']['next_results']
    except KeyError, e:
        break
    kwargs = dict([kv.split('=') for kv in next_results[1:].split("&")])
    search_results = twitter_api.search.tweets(**kwargs)
    statuses += search_results['statuses']
    screen_names = [user_mention['screen_name']
                    for status in statuses
                    for user_mention in status['entities']['user_mentions']]
    screen_names=set(screen_names)
    screen_names=list(screen_names)
    ans=1;
    while(ans==1):
        os.system('clear')
        ch=int(input('1) Small World Network for Users\n2) Random Network for Users\n3)
Small World Network for Retweeters\n4) Random Network for Retweeters\nEnter choice: '));
        if(ch==1):
            plt.subplot(1,2,1)
            G=nx.watts_strogatz_graph(8, 2, 0.7);
            plt.title('Small World Network for Users');
            nx.draw(G, with_labels = True);
            print('\nSmall World Network for Users')
            print('\nDegree Centrality')
            t1=nx.degree_centrality(G)
            print(nx.degree_centrality(G));
            print('\nMost Important node based on Degree Centrality: ')
            v1=[]
            for i in t1.keys():
                if(t1[i]==max(t1.values())):
                    v1.append(i);
                    print('Node '+str(i));
            print('\nCloseness Centrality')
            t2=nx.closeness_centrality(G)
            print(nx.closeness_centrality(G));
            print('\nMost Important node based on Closeness Centrality: ')
            v2=[]
            for i in t2.keys():
                if(t2[i]==max(t2.values())):
                    v2.append(i);
                    print('Node '+str(i));
            print('\nBetweenness Centrality')
            t3=nx.betweenness_centrality(G)
            print(nx.betweenness_centrality(G));
            print('\nMost Important node based on Betweenness Centrality: ')
            v3=[]
            for i in t3.keys():
                if(t3[i]==max(t3.values())):
                    v3.append(i);
                    print('Node '+str(i));
            print('\n\nMost Important Node in the network: ')
            for i in v1:
                if(i in v2 and i in v3):
                    print('Node '+str(i)+' - '+screen_names[i]);
            plt.subplot(1,2,2)
            G=nx.Graph();
            nx.draw(G);
            for i in range(8):
                plt.text(0.5, 0.2+(i*0.08), str(i)+'-'+screen_names[i],
horizontalalignment='center', verticalalignment='center', fontsize=9)
            plt.savefig('social1/'+q+'1.png')
            plt.show();
            plt.close('all');
        elif(ch==2):
            plt.subplot(1,2,1)
            G=nx.gnp_random_graph(8,0.2);
            plt.title('Random Network for Users');
            nx.draw(G, with_labels = True);

```

```

print('\nRandom Network for Users')
print('\nDegree Centrality')
t1=nx.degree_centrality(G)
print(nx.degree_centrality(G));
print('\nMost Important node based on Degree Centrality: ')
v1=[]
for i in t1.keys():
    if(t1[i]==max(t1.values())):
        v1.append(i);
        print('Node '+str(i));
print('\nCloseness Centrality')
t2=nx.closeness_centrality(G)
print(nx.closeness_centrality(G));
print('\nMost Important node based on Closeness Centrality: ')
v2=[]
for i in t2.keys():
    if(t2[i]==max(t2.values())):
        v2.append(i);
        print('Node '+str(i));
print('\nBetweenness Centrality')
t3=nx.betweenness_centrality(G)
print(nx.betweenness_centrality(G));
print('\nMost Important node based on Betweenness Centrality: ')
v3=[]
for i in t3.keys():
    if(t3[i]==max(t3.values())):
        v3.append(i);
        print('Node '+str(i));
print('\n\nMost Important Node in the network: ')
for i in v1:
    if(i in v2 and i in v3):
        print('Node '+str(i)+' - '+screen_names[i]);
plt.subplot(1,2,2)
G=nx.Graph();
nx.draw(G);
for i in range(8):
    plt.text(0.5, 0.2+(i*0.08), str(i)+'-'+screen_names[i],
horizontalalignment='center', verticalalignment='center', fontsize=9)
plt.savefig('social1/'+q+'2.png')
plt.show();
plt.close('all');
elif(ch==3):
plt.subplot(1,2,1)
G=nx.watts_strogatz_graph(8, 2, 0.8);
plt.title('Small World Network for Retweeters');
nx.draw(G, with_labels = True);
print('\nSmall World Network for Retweeters')
print('\nDegree Centrality')
t1=nx.degree_centrality(G)
print(nx.degree_centrality(G));
print('\nMost Important node based on Degree Centrality: ')
v1=[]
for i in t1.keys():
    if(t1[i]==max(t1.values())):
        v1.append(i);
        print('Node '+str(i));
print('\nCloseness Centrality')
t2=nx.closeness_centrality(G)
print(nx.closeness_centrality(G));
print('\nMost Important node based on Closeness Centrality: ')
v2=[]
for i in t2.keys():
    if(t2[i]==max(t2.values())):
        v2.append(i);
        print('Node '+str(i));
print('\nBetweenness Centrality')
t3=nx.betweenness_centrality(G)
print(nx.betweenness_centrality(G));
print('\nMost Important node based on Betweenness Centrality: ')
v3=[]

```



```

        for i in t3.keys():
            if(t3[i]==max(t3.values())):
                v3.append(i);
                print('Node '+str(i));
        print('\n\nMost Important Node in the network: ')
        for i in v1:
            if(i in v2 and i in v3):
                print('Node '+str(i)+' - '+screen_names[i]);
    plt.subplot(1,2,2)
    G=nx.Graph();
    nx.draw(G);
    for i in range(8):
        plt.text(0.5, 0.2+(i*0.08), str(i)+'-'+screen_names[i],
horizontalalignment='center', verticalalignment='center', fontsize=9)
    plt.savefig('social1/'+q+'.png')
    plt.show();
    plt.close('all');
else:
    plt.subplot(1,2,1)
    G=nx.gnp_random_graph(8,0.2);
    plt.title('Random Network for Retweeters');
    nx.draw(G, with_labels = True);
    print('\nRandom Network for Retweeters')
    print('\nDegree Centrality')
    t1=nx.degree_centrality(G)
    print(nx.degree_centrality(G));
    print('\nMost Important node based on Degree Centrality: ')
    v1=[]
    for i in t1.keys():
        if(t1[i]==max(t1.values())):
            v1.append(i);
            print('Node '+str(i));
    print('\nCloseness Centrality')
    t2=nx.closeness_centrality(G)
    print(nx.closeness_centrality(G));
    print('\nMost Important node based on Closeness Centrality: ')
    v2=[]
    for i in t2.keys():
        if(t2[i]==max(t2.values())):
            v2.append(i);
            print('Node '+str(i));
    print('\nBetweenness Centrality')
    t3=nx.betweenness_centrality(G)
    print(nx.betweenness_centrality(G));
    print('\nMost Important node based on Betweenness Centrality: ')
    v3=[]
    for i in t3.keys():
        if(t3[i]==max(t3.values())):
            v3.append(i);
            print('Node '+str(i));
    print('\n\nMost Important Node in the network: ')
    for i in v1:
        if(i in v2 and i in v3):
            print('Node '+str(i)+' - '+screen_names[i]);
    plt.subplot(1,2,2)
    G=nx.Graph();
    nx.draw(G);
    for i in range(8):
        plt.text(0.5, 0.2+(i*0.08), str(i)+'-'+screen_names[i],
horizontalalignment='center', verticalalignment='center', fontsize=9)
    plt.savefig('social1/'+q+'.png')
    plt.show();
    plt.close('all');
    ans=int(input('\nDo you want to continue?'))

def advance_analysis():
    CONSUMER_KEY = '2D2qQNR2rSUwzulwEGFPXR3zp'
    CONSUMER_SECRET = 's7A2pwwTvo0dZvzbA5lzP06YG9tm3IST9qZjkAkF2spKrXxyW1'
    OAUTH_TOKEN = '1007573251231383553-nu10kcCDaJDLd9UDVsVZx6FFKy5BKo'
    OAUTH_TOKEN_SECRET = '2L10gEAkQnGeZe4g88d13mgywW6e0oRYkiBN8kFhvxIE1'

```

```

auth = twitter.oauth.OAuth(OAUTH_TOKEN, OAUTH_TOKEN_SECRET, CONSUMER_KEY,
CONSUMER_SECRET)
twitter_api = twitter.Twitter(auth=auth)
q = str(input('Enter a Name to Search: '))
count = int(input('\nEnter the number of tweets to fetch: '))
search_results = twitter_api.search.tweets(q=q, count=count)
statuses = search_results['statuses']

for _ in range(5):
    try:
        next_results = search_results['search_metadata']['next_results']
    except KeyError, e:
        break
    kwargs = dict([kv.split('=') for kv in next_results[1:].split("&")])
    search_results = twitter_api.search.tweets(**kwargs)
    statuses += search_results['statuses']

status_texts = [status['text']
                 for status in statuses]

screen_names = [user_mention['screen_name']
                 for status in statuses
                 for user_mention in status['entities']['user_mentions']]

hashtags = [hashtag['text']
             for status in statuses
             for hashtag in status['entities']['hashtags']]

words = [w
          for t in status_texts
          for w in t.split()]

screen_names=set(screen_names)
screen_names=list(screen_names)
print(' \n\n Texts From top 5 Tweets : ')
print(json.dumps(status_texts[0:5], indent=1))
print(' \n\n User Names For top 5 Tweets : ')
print(json.dumps(screen_names[0:5], indent=1))
print(' \n\n Hashtags From top 5 Tweets : ')
print(json.dumps(hashtags[0:5], indent=1))
print(' \n\n Collection of the words collected from the tweets : ')
print(json.dumps(words[0:5], indent=1))
retweets = [
    (status['retweet_count'],
     status['retweeted_status']['user']['screen_name'],
     status['text'])

    for status in statuses

    if status.has_key('retweeted_status')
]

print(' \n\n Top Users who use this keyword: ')
print(json.dumps(screen_names[0:8], indent=1))

print(' \n\n Top Retweeters for the Tweet : ')
retweets1 = twitter_api.statuses.retweets(id=317127304981667841)
print([str(r['user']['screen_name']) for r in retweets1])

for item in [words, screen_names, hashtags]:
    c = Counter(item)

def lexical_diversity(tokens):
    try:
        return 1.0 * len(set(tokens)) / len(tokens)
    except:
        return 0;

def average_words(statuses):
    total_words = sum([len(s.split()) for s in statuses])

```

```

        try:
            return 1.0 * total_words / len(statuses)
        except:
            return 0;

print('\n\n Lexical Diversity in Words : ')
print lexical_diversity(words)
print('\n\n Lexical Diversity in Screen Names : ')
print lexical_diversity(screen_names)
print('\n\n Lexical Diversity in Hashtags : ')
print lexical_diversity(hashtags)
print('\n\n Lexical Diversity in Status Texts : ')
print average_words(status_texts)

sys.stdin.read(1)

def analyze_tweets():
    api = TwitterClient()
    q = input('Enter a Name to Search: ')
    tweets = api.get_tweets(query=q, count=1000)
    ptweets = [tweet for tweet in tweets if tweet['sentiment'] == 'positive']
    print("\n\nGood tweets percentage: {} %".format(100 * len(ptweets) / len(tweets)))
    ntweets = [tweet for tweet in tweets if tweet['sentiment'] == 'negative']
    print("\n\nBad tweets percentage: {} %".format(100 * len(ntweets) / len(tweets)))
    print("\n\nUnbiased tweets percentage: {} %".format(100 * (len(tweets) -
len(ntweets) - len(ptweets)) / len(tweets)))
    t1 = Texttable()
    t2 = Texttable()
    print "\n\n"
    t1.add_row(['Good Tweets']);
    for tweet in ptweets[:10]:
        t1.add_row([tweet['text']]);
    print t1.draw()
    print "\n\n"
    t2.add_row(['Bad Tweets']);
    for tweet in ntweets[:10]:
        t2.add_row([tweet['text']]);
    print t2.draw()
    sys.stdin.read(1)

def search_tweet():
    api = TwitterClient()
    t = Texttable()
    print('\n1) Search by keyword\n2) Search by username');
    choice=int(input('Enter choice: '));
    if(choice==1):
        q = str(input('Enter a Name to Search: '))
        nu = int(input('\nEnter the number of tweets to fetch: '))
        tweets = api.get_tweets(query=q, count=100)

        i=1;
        t.add_row(['', 'Tweets']);
        for tweet in tweets[:nu]:
            t.add_row([i, tweet['text']]);
            i+=1;
        print t.draw()
    elif(choice==2):
        extractor = twitter_setup()
        q = str(input('Enter a Name to Search: '))
        nu = int(input('\nEnter the number of tweets to fetch: '))
        try:
            tweets = extractor.user_timeline(screen_name=q, count=100)
            t.add_row(['', 'Tweets']);
            i=1;
            for tweet in tweets[:nu]:
                t.add_row([i, tweet.text]);
                i+=1;
            print t.draw()
        except:

```

```

        print('User not found');
    sys.stdin.read(1)

def user():
    ch = 1
    while ch != 7:
        os.system('clear')
        print("-" * 50)
        print("User Menu")
        print("-" * 50)
        print('\n1) View Trends \n2) Search Tweets \n3) Analyze Tweets \n4) Advanced
Tweet Analysis \n5) Community Detection and Graph Creation\n6) Network Models and
Computation \n7) Exit')
        ch = int(input('\nEnter your choice : '))
        if ch == 1:
            view_trends()
        elif ch == 2:
            search_tweet()
        elif ch == 3:
            analyze_tweets()
        elif ch == 4:
            advance_analysis()
        elif ch == 5:
            combined_graph()
        elif ch == 6:
            network_models()
        elif ch == 7:
            print('\nApplication Closing')
        else:
            print('\nWrong choice entered, Please try again')
            sys.stdin.read(1)

def main():
    user()

main()

```