

# **Pandemic : The Untold Experiences**

## **Introduction**

Pandemic witnessed a change in our norms around work discipline, educational institutions, healthcare systems and a number of other institutions had changed but unfortunately, the understanding of gender and sexuality still remained the same. Our traditional understanding of gender identities and sexualities excludes a section of our society and as a consequence makes them invisible. Furthermore, this understanding of gender and sexuality made pandemic a whole lot difficult for communities who do not identify as either 'man' or a 'woman' (Shah et al., 2015) This is why a project like this is essential where we understand the lived realities of the marginalized communities during the pandemic.

During the pandemic, due to the lockdowns and the nationwide restrictions of traveling, the only convenient medium of communication remained social media. Social media platforms are a great medium through which people stay in touch with their loved ones, form new connections by meeting new people, share their life stories, find people with common interests, and more. In short, these spaces bring the world a little closer and make it easier for people to express themselves and share their stories with everyone else.

Just like everyone else, communities marginalized on the basis of their gender identity and sexuality also took their issues online and more so during the pandemic. Social media platforms like Twitter, Reddit, Instagram and Facebook have been the major platforms where experiences of everyone during the pandemic can be found. Therefore, it becomes necessary for us to use the data available on these platforms to suitably understand the experiences of others.

Therefore, for a project aimed at understanding the lived realities of the marginalized communities during the pandemic, it is essential to use data available on these social media platforms and our project seeks to do just that.

To sum it up, the purpose of our project is to understand the experiences of the communities marginalized on the basis of gender identities and sexualities. It also seeks to recognize how effective computational tools are in the study of gender in the context of marginalized communities.

## Literature Review

The Government of India imposed a nationwide lockdown on 24th March, 2020 for 21 days. This lockdown restricted the movement of 1.38 billion residents. Furthermore, offices, schools and colleges shut down and students were sent back home which resulted in online classes, work from home and almost everything became virtual.

There was also a temporary suspension of flights from and to India which led to situations where a lot of us could not go back home due to the travel restrictions related to COVID-19 which, for a lot of us, meant staying with our relatives and other friends for almost 4 months. Travel restrictions worst affected migrant workers who are essentially the backbone of India's urban economy. They trekked hundreds of kilometers to reach their villages during lockdown without proper availability of transport (Salihefendic, 2016). Thus, the suspension of all modes of transport affected the entire population on varying levels.

During the pandemic, there was also shortage of groceries, more challenges to the already overstretched public health system of India, new COVID protocols which included wearing masks, maintaining physical distance of at least 2metres, using sanitizers along with numerous other challenges.

With all these challenges, came some realizations. All of us realized the importance of touch, affection, communication and most importantly, family and friends. Lucky are the ones who have someone to call family, family that supports them in tough times like the pandemic. But the pandemic also made us realize that injustice does not affect everybody equally, the pandemic did not affect everyone equally (Roy, 2020).

Pandemic witnessed a change in our norms, a reconstruction of our norms. Norms around work discipline, educational institutions, healthcare systems and a number of other institutions had changed but unfortunately, the understanding of gender and sexuality still remained the same. When it comes to gender identities, all basic rights in India have been easily accessible to only two genders namely, 'men' and 'women'. Our traditional understanding of gender and sexuality made pandemic a whole lot difficult for communities who do not identify as either 'man' or a 'woman' (Shah et al., 2015). This traditional understanding of gender identities and sexualities excludes a section of our society and as a consequence makes them invisible. Therefore, there is an urgent need to shift from the traditional ideas to a more inclusive definition of gender and sexuality (Shah et al., 2015, p. 13).

One community that does not conform to the norms around gender binary that I think is particularly important to study is the transgender community, thus this study's primary subjects belong to the transgender community.. Before we look into the research done on this particular group, I will first clarify the reason for my interest in this community.

'Transgender' refers to people who "transgress the rigid gender dichotomy" (Ganguly & Singh,

2021). The community is not as simple to understand as this definition, their community comes with their own complexities and cultural diversities just like any other community (Ganguly & Singh, 2021). Trans people are more likely to be unemployed since they have low chances of being accepted in a traditional job. According to a report by the National Human Rights Commission, there are 92% of transgender communities who are deprived of the right to participate in economic activities and 96% are forced to take up jobs that are low paying which includes sex work, begging and other activities which are exploitative in nature. 23% of transgender communities participate in sex work itself. Moreover, 29% transgenders have never attended school while only around 30% have graduated with a degree (Ganguly & Singh, 2021).

The transgender community always had to fight for the most basic rights given to everyone since birth. The above statistics act as evidence to this. They only truly got freedom of expression after the Supreme Court verdict in April 2014. The verdict states that “Transgender will be included as a third gender under the various scholarship/fellowships schemes of UGC” (Shah et al., 2015). This judgement gave recognition to the marginalized communities and acknowledged individuals who self-identified as neither a ‘man’ or a ‘woman’ but the verdict was vague in itself and had a number of unanswered questions, questions like, who exactly is transgender? This verdict was also incapable of supporting communities marginalized on the basis of their gender expression (Shah et al., 2015, p. 12). Even though the April 2014 verdict gave a means to protect the basic rights of the marginalized communities, the general public still does not accept everyone who does not place themselves in the gender binary.

The above mentioned statistics are from the pre-pandemic period. It is necessary to figure out how the pandemic became a significant factor in changing these statistics and whether the pandemic made things worse or better.

The pandemic changed everyone’s lives. We all have stories from the pandemic but it is also important to be aware of others’ experiences. While some of us were home fighting loneliness from not meeting our friends for a long time, there were many others who were fighting to get one meal a day. In other words, our respective privileges determined our experiences during the pandemic and arguably some people suffered more than others. Millions in our country lost their jobs ergo a stable livelihood, but the transgender community who face discrimination even in a “normal” setting, the pandemic only worsened their state of living, as can be inferred from the above mentioned statistics and experiences.

Ahmed and Sifat conducted a rapid perception survey in 2021 amongst the transgender people in Bangladesh to understand the impact of COVID-19 on the transgender community (Ahmed & Sifat, 2021) and found that 18% of the participants living in urban areas and 10% of the participants living in rural areas had no food stored at home. The lockdowns being extended along with the stress of pandemic only added to their hopelessness. It was also found that 61% of the respondents needed food, 59% needed money, 9% needed secure work, 9% needed medicine

and 8% needed protective equipment. They also experienced financial difficulties due to the lack in social protection, this lack in social protection can be attributed to the fact that the transgender people are not publicly identified (Ahmed & Sifat, 2021).

Scholars agree that gender dysphoria generally ends up putting the transgender community at a higher risk of anxiety, depression, self-harm and suicidal tendencies (Roy, 2020). They often face physical abuse and find support in NGOs or other community-based organizations aimed to provide such marginalized communities shelter and other help (Roy, 2020). The lived reality of the transgender community is facing social stigma on the basis of their identity and being denied fundamental human rights and basic privileges. Furthermore, the transgender people are exploited by the property owners, they face financial insecurity thus there is a never ending struggle for finding a safe home and at least two meals a day. Due to the widespread transphobia and lack of support they are forced to take up jobs where they get exploited. They take up jobs like sex work, begging and seeking alms in functions. Noticeably, the transgender community always faces twice the discrimination due to their gender and their socio-economic status (Ganguly & Singh, 2021).

One of the main concerns during the pandemic was health care services for the transgender community. The healthcare services had to be expanded to accommodate the ever increasing number of covid positive patients. But for transgender people, even seeking medical help during the pandemic was a task. Healthcare practitioners would not know which ward ('male' or 'female') to place the trans person in. Placing them in either wards would make the other patients uncomfortable. Sometimes, the healthcare practitioners themselves were transphobic and did nothing to help the trans person. In addition to this, many trans people were undergoing the sex reassignment surgeries (SRS) but with the pandemic, all such surgeries had to be postponed indefinitely since there were no doctors available in the outpatient departments (Ganguly & Singh, 2021 ; Roy, 2020). This resulted in more frustration, confusion and depression in transgender communities.

To tackle this issue of discrimination faced in healthcare institutions, there have been various discussions about the necessity to make the Indian healthcare system more inclusive for trans people (Pandya & Redcay, 2021). Both the public and the private sectors of the healthcare system need to be trans-friendly and have to be trained to understand the healthcare needs of trans people. In a survey to obtain transgender community response on the awareness around the pandemic (Ahmed & Sifat, 2021), it was found that while all transgender communities knew about coronavirus, only 8.8% shield themselves while coughing and sneezing to prevent the spread, 47.5% wore masks and only 21.3% used sanitizers. The given statistics prove that these communities have been neglected and forgotten by healthcare and other stakeholders responsible for healthcare. But healthcare systems are not the only institutions where trans people faced discrimination during the pandemic crisis.

Scholars have done a lot of research on the life of the transgender communities to understand the lived realities of trans people during the pandemic. A few studies have tried to categorize the results they obtained into themes (Pandya & Redcay, 2021 ; Bhalla & Agarwal, 2021). These themes captured and categorized the lived experiences of the participants. One of the papers (Bhalla & Agarwal, 2021) which used themes to describe the lived experiences selected participants through snowball sampling. The study was qualitative in nature and applied transcendental phenomenology as a mode of enquiry whereas another paper (Pandya & Redcay, 2021) which used themes categorized the questions for their telephonic interviews into five categories. They used a qualitative descriptive method which helped them develop descriptions of their participants' experiences during the lockdown. The participants in this study as well were selected through snowball sampling.

I noticed an overlap in the methods used by these papers. Both the above mentioned papers used snowball sampling to select their participants and then proceeded to hold telephonic interviews. One reason I can think of for both the studies using snowball sampling even though snowball sampling usually leads to biased results is that communities marginalized on the basis of gender identities and sexualities usually exist in groups and communities. Other marginalized individuals who are not a part of such groups and communities either are hard to get in touch with or are not ready to participate in such studies. These studies also had an approach which was qualitative in nature. Almost all papers that I have come across have used a qualitative methodology in addition to some other approach. These other approaches ranged from interpretive approaches, content analysis (Ahmed & Sifat, 2021) to using an intersectionality lens that provides a new perspective on the interconnectedness between various social categorisations and through this lens understanding how power and inequality operates and impacts traditionally oppressed groups (Bhalla & Agarwal, 2021).

One paper in particular that interested me the most conducted a cross-sectional online survey of LGBTQ+ people to collect data on their experiences in the pandemic and conduct regression analyses for a range of sociodemographic variables (Kneale & Bécares, 2021). This was one of the few papers that I had come across and had used a quantitative methodology to form conclusions.

The work done by Kneale and Bécares tried to assess the mental health and experiences of discrimination among the LGBTQ+ people at the start of the COVID-19 pandemic. They collected data from a cross-sectional online survey conducted in April-July of 2020 where they asked questions related to mental health, experiences of discrimination and other pandemic-related experiences. They used regression analysis on a range of sociodemographic variables to understand the relationship between sexual orientation and gender and mental health and experiences of discrimination. This study found that the majority of their participants had

high levels of depressive symptoms. In addition to that, the LGBTQ+ community experienced high levels of stress, depression and discrimination during the COVID-19 pandemic.

Through this extensive reading on the pandemic experiences of the transgender community, it is clear that there has been a lot of studies and papers written on the pandemic experiences of the communities marginalized on the basis of gender identities and sexualities in India. Most of them have used a qualitative method to form conclusions with some interviewing and learning about the different stories and experiences in-depth, while others reading articles, papers and journals published on PubMed, PsycINFO, Medline and Google Scholar to write an extensive and elaborate literature review (Philip, 2021). These qualitative studies have given us enough evidence that prove that pre-existing stigma, inequalities and violence against these communities had intensified during the COVID-19 pandemic. Some of the quantitative studies we came across also used data collected through a qualitative approach and made inferences on the basis of that. But one very important means of communication that was significant during the pandemic is never mentioned by any of the studies described above, that is, communication through social media.

Social media platforms reveal information that might not be available in data obtained from other sources. These platforms have their own norms, their own way of expressing things and the data obtained from such platforms provide insights which arguably cannot be found elsewhere. In addition to this, they are a great medium through which people stay in touch with their loved ones, form new connections by meeting new people, share their life stories, find people with common interests, in short, these spaces bring the world a little closer and makes it easier for people to express themselves and share their stories with everyone else.

Just like everyone else, communities marginalized on the basis of their gender identity or expression also took their issues online and more so during the pandemic. Social media platforms like Twitter, Reddit, Instagram and Facebook have been the major platforms where experiences of many people during the pandemic can be found. Consequently, crucial debates and discussions from the pandemic crisis have been developed on these platforms making them a great primary source to study the experiences of transgender during the pandemic.

All the above mentioned studies used semi-structured interviews, structured interviews or surveys to collect their data during the lockdown or post-lockdown. As mentioned, social media was a means through which everyone stayed connected and a place where they asked for help due to its great reach. By using data available on social media, it becomes easier to understand the experiences of a number of people not situated in one particular geographical location. It helps us capture the experiences of trans people differently than what has been done till now. Using this social media data as my primary data source, I would then proceed to conduct regression analysis where my variables would be the time period and other socio-demographic variables. By using a computational tool like regression analyses, the experiences of the trans

community can be captured differently offering a new perspective and can add to the existing knowledge collection on the trans people pandemic experiences.

This project will be focused on answering the following questions :

- To understand what are the most widely shared pandemic experiences of transgender people in India on social media
- To assess and analyse what the realities of transgender people have been during the pandemic, and what that reveals about their place in the Indian polity
- To gauge how effective the use of computational tools is in the study of gender, particularly in the context of marginalized communities

Through computational analysis of social media data about transgender pandemic experiences, we will receive insights which are unique and different from what has been observed so far. I will also relate and interpret my findings and results from the computational analysis thus doing a solid analysis and reaching legitimate conclusions.

## Method Description

For this project, data from the ‘lgbt’ subreddit was analyzed. The data obtained from this subreddit mainly consisted of posts and the title of the post. In Reddit, the posts are of three types, namely, videos, images and text. Users share links, images, and/or text on a subreddit of their choice and thus interact through posts and comments. Subreddits are forums/groups on Reddit and are created for a specific purpose. For our project, we used the ‘lgbt’ subreddit and we used two methods for obtaining our dataset and further analyzing it, namely, through Reddit API and through Pushshift API.

### Reddit API

By using Reddit API, we were able to obtain data of hot posts. Hot posts are a type of post on reddit which are classified as given the most upvotes during a particular day. Reddit scores its posts on the basis of upvotes, downvotes and the age of the post. The older the post gets, the smaller the score becomes (Salihefendic, 2016).

Therefore, posts that are classified as hot posts are no older than a day. For our research, we chose the hot posts from 13th May, 2022. The reason for choosing this particular date was because of the following :

1. Since, 13th May was a Saturday, a weekend, we expected there to be greater interaction among the users of the subreddit.
2. One significant drawback of the Reddit API which has been explained in detail below.

One major drawback of Reddit API, which also became our cause for implementing Pushshift API is its inability to provide us with a way to only observe posts from a particular time period. We essentially wanted to observe the behavior of users from lgbt subreddit during the pandemic days, but due to the inability to obtain time specific posts, we decided to analyze data which was most popular among the users hence the hot posts from a Saturday.

Thus, using this method we obtained data from the lgbt subreddit, a data that showcased posts that are posted in this subreddit on an average day thus giving us a general view.

### Pushshift API

It was while we obtained data through Reddit API, we came across Pushshift API which also has the ‘after’, ‘before’ parameter which lets us input a date as a starting point and as an ending point respectively. For our research, we are looking at the experiences of the lgbtq+ community during pandemic, thus our after timestamp was 10th March, 2019 and the before timestamp was 31st August, 2021. During March, 2019 India had imposed travel restrictions along with suspending visas since the World Health Organization had declared COVID-19 a pandemic with around 114 countries fighting the virus. It was during June 2020, it came to everyone’s attention that around



a 100 million daily wage workers had lost their jobs and these migrant workers were the worst affected from the pandemic in terms of job opportunity and security. By August, 2020 countries had discovered and developed methods to test the virus such as the RT-PCR along with announcing vaccines. Many countries had already started making it mandatory for the residents to take 2 doses of vaccine. It also included a 'keyword' parameter which helped us choose a word that we needed to be present in the posts posted during pandemic. For the purpose of our research, the chosen keyword was 'pandemic'.

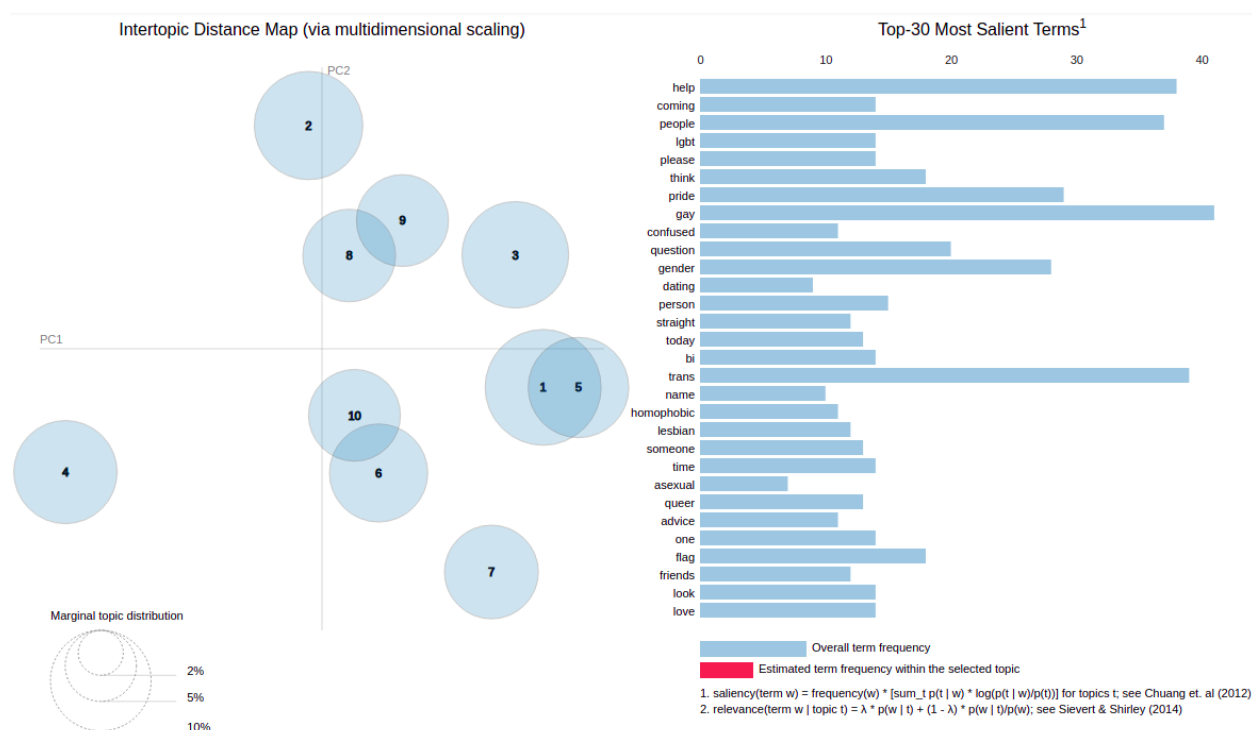
Through this method, we obtained the title of all the posts from the time period mentioned above. One aspect which we could have improved in this method is using the content of all the posts instead of just the title of the post but we could not find a parameter in this API that would help us accomplish the same.

Using this method, we were able to obtain titles of posts from the 'lgbt' subreddit that had a very pandemic specific view and it provided us with keywords that helped us in understanding the idea of the post. Our assumption here is that the titles of the posts give a gist and an insight to the content of the post itself and that there is connection between the posts and their titles.

Initially, we decided to scrape five subreddits during the time period of the pandemic. But after receiving the data, we realized three of the five subreddits we have chosen did not have enough users to give us enough data. The remaining two subreddits gave us enough data to work with but by using two separate subreddits, the chances of variety in data increases which could have given us results with varying topics. Due to this reason, we decided to use lgbt subreddit instead of using lgbt and TransgenderIndia subreddit both where one includes various other communities and the other is specific to transgenders.

After using the above two methods, the data received was cleaned, that is, pre-processing the data, and other aspects of topic modeling like tokenizing the data obtained, preparing it for the final modeling, etc was performed. To achieve the model, Latent Dirichlet Allocation (LDA) was performed. The LDA topic modeling infers the possible topics present in the text based on words from the text and the inference is based on a Bayesian framework. It uses a generative probabilistic model and Dirichlet distributions to discover the abstract topics present in the text given for processing (Giri, 2022).

The text scraped from the 'lgbt' subreddit was divided into 10 topics and each topic was analyzed for the top-30 most relevant terms. In addition to this, the models created provided an estimate of the overall term frequency in the entire text provided for modeling and the estimated term frequency within the selected topic. An instance of the model has been shown below to understand how to interpret the output of the model.



The circles on the left indicate the 10 topics and the bars on the right are the top-30 most salient terms considering all the topics. The figure on the left is a visualization of the topics in a 2-D space and the area of these topic circles is proportional to the amount of words (present on the right) that are a part of these topics (F, 2021).

By clicking on either of the circles, the bars on the right represent the top-30 most relevant terms. The ‘Analysis and Findings’ section uses the relevant terms from analyzing the topics.

The red bars indicate the estimated term frequency of the words within the selected topic and the blue bars indicate the overall term frequency.

The final models obtained for both the methods and our analysis of them have been explained in detail in the ‘Analysis and Findings’ section.

## Analysis and Findings

Reddit was founded in 2005 and was meant for users to share content such as links, posts etc. which can be of three types, namely, images, videos, and articles. These submissions are upvoted by other users and on the basis of these upvotes, the submissions are further classified into hot posts, top posts, and best posts. Reddit is made up of numerous forums and groups called Subreddits. For the purpose of our research, we will be working mainly with the 'lgbt' subreddit.

The 'lgbt' subreddit was founded in 2008 and according to the authors, the following forum was created as “A safe space for GSRM (Gender, Sexual, and Romantic Minority) folk to discuss their lives, and issues, interests, and passions. LGBT is still a popular term used to discuss gender and sexual minorities, but all GSRM are welcome beyond lesbian, gay, bisexual, and transgender people who consent to participate in a safe space.” (Rose, 2008)

Since the research's main focus is on the lives of the communities marginalized on the basis of gender and sexual minorities, this platform provided us with the appropriate data for our research.

As mentioned in the methodology section, we used the Reddit API to recognize the topics that were discussed on any random day. Through this, we found that the words used majorly included gay, trans, lesbian, gender, question, think, lgbtq, feel, and help. Other words that could be found in a few posts included sexuality, queer, bisexual, bi, and asexual. Only words that are significant for our analysis have been mentioned as examples. Refer to the word cloud attached below to find out all the words that were found in the hot posts on the 13th of May.



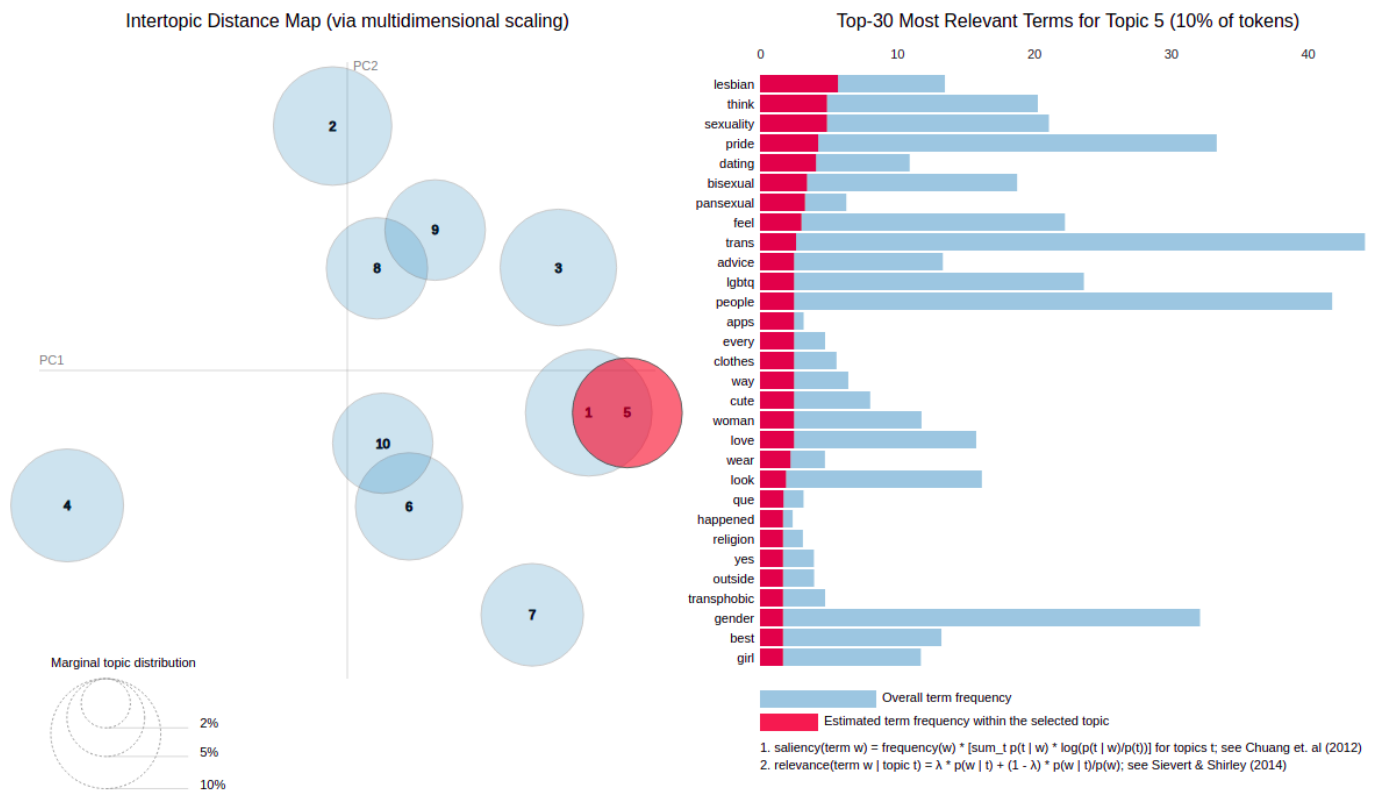
For the topic modeling performed on the data retrieved from the hot posts, the entire text was divided into 10 topics and each topic was reviewed and analyzed for the top-30 most relevant terms. The most commonly used words in the 10 topics were gay, help, trans, people, lesbian, help, think, LGBT, coming, and friends. Other frequently occurring words included confused, pride, gender, protest, religion, bisexual, happy, and homophobic. Since, the entire purpose of the

LGBT subreddit was for the GSRM community to discuss their interests, experiences, and passions, the fact that words like gay, trans, lesbian, and lgbt are commonly used would make sense. After analyzing some of these posts manually, we infer that these words have been used by the users to introduce themselves, share their experiences with other users who also have the same gender and/or sexuality, and tell them about their experiences as a person who was a part of LGBTQ+.

The terms such as pride, queer, gay, lesbian, trans, LGBT, and bisexual have been used in posts that focus especially on the queer community. Some posts signify the importance of freedom and that they will keep fighting and ‘protesting’ till all people have human rights. One of the posts said, “My super queer tattoo says Freedom and basically means no rest until all people have human rights.” These terms are noticed in posts extending support and love to everyone else struggling and also as a motivation to continue fighting for human rights. One post read, “I have a family reunion this summer and a few cousins are openly homophobic. Today I found a shirt to wear to support my one gay cousin.” The shirt read “Be gay, do crimes”.

Therefore, we notice that users have built a network and a support system where they share the latest updates, protest details, support and motivate each other and act as a unit against the discrimination faced by them.

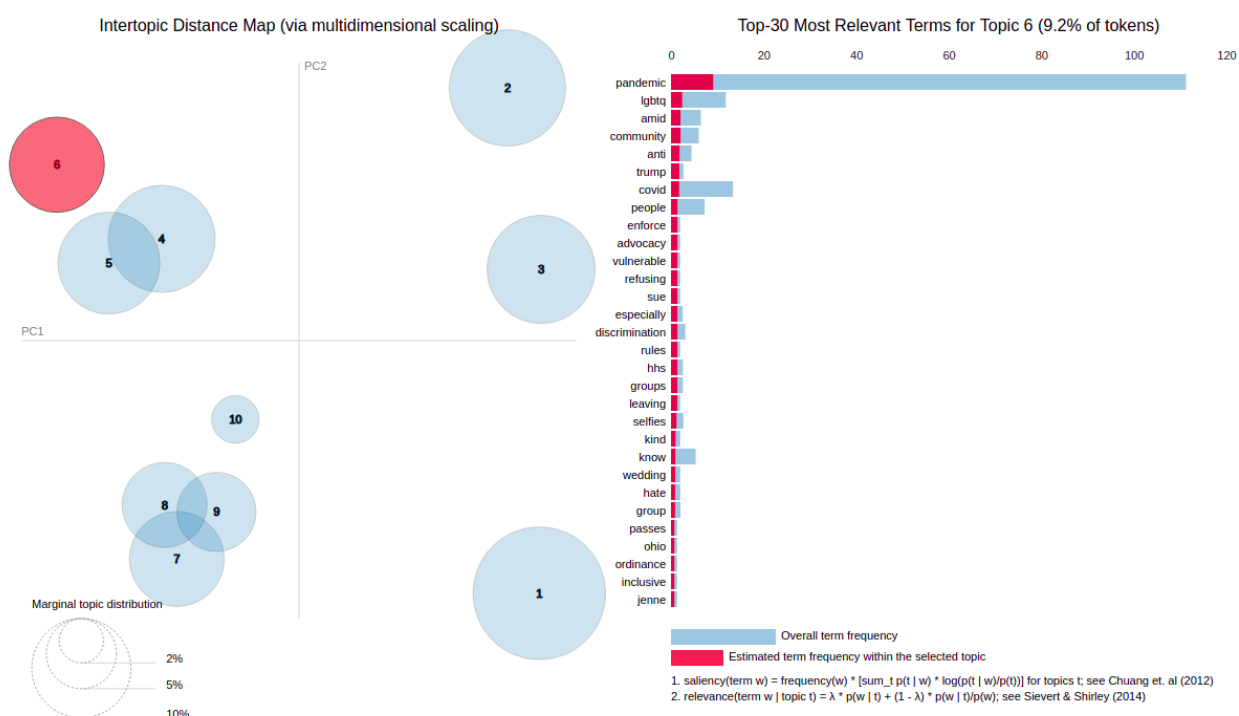
Following is the result obtained from topic 5 through modeling :





The term ‘surgery’ was indicative of the Sex Reassignment Surgery (SRS) which could not be performed during the pandemic. Stopping the surgery in between led to hormonal imbalance and more anxiety during this period of isolation (Stevens & Nikos-Rose, 2021). One of the posts read, “As for many others, my finances got wrecked due to the pandemic and my long awaited surgery had to be postponed to an unknown date. If you could please share this post it would mean the world to me.” The post contained a link for providing funds.

Following is the result obtained from topic 6 through modeling :



The term pandemic is the most frequently occurring in all the topics since the keyword chosen by us as a parameter was pandemic. Thus, the occurrence of the word pandemic as a relevant term was expected. Terms like vulnerable, stress, friends, health, surgery, aids and other similar words might have appeared due to the issues faced by the communities during the pandemic. For instance, the Trump administration had reversed the protections put in place by Barack Obama in health care based on gender identities during the pandemic encouraging health care providers to deny care to transgender people. This is an example from the states, there have been many instances from India itself, which proves that communities marginalized on the basis of sexuality and gender identities had more difficulties during pandemic than other cis men and women.

This particular method only used the title of the posts and saved them in a csv file making it easy for us to access and review it for ourselves. Thus, the following section is an account of our understanding of the topics and the time period they are based in.

Many of the titles talk about their struggle to deal with the pandemic. The users have shared links to discovering how their communities were treated in the previous pandemics such as the HIV/AIDS pandemic which began in 1980 and had peaked in 2005-2006, the ebola epidemic from 2013 to 2016, and how had the non-Queer community reacted to the discrimination faced by the lgbtq+ communities. The users had also shared links regarding mental health and how it was affected during the pandemic. One significant point made by numerous users was the comparison of the COVID-19 pandemic with HIV/AIDS. These included ways to learn to survive the pandemic, what was the response of the non-Queer community to these two pandemics and more.

The HIV epidemic disproportionately affected certain members of the LGBTQ+ community. The LGBTQ+ community has faced discrimination in many aspects of their life including housing, public accommodations, and more but most importantly in the health care institutions. They were denied basic health care facilities due to the stigma and discrimination surrounding them. They were also prone to taking up more jobs such as sex work and were thus hit especially hard by the epidemic. The reason for comparing the HIV epidemic and the COVID-19 pandemic could be attributed to the discrimination that is still faced by them and the fact that this time it was not just the LGBTQ+ community that was hit hard by the virus but also the non-queer community.

Our research had initially decided to focus only on the transgender community in India but due to insufficient data on Reddit, we had to expand our research to the lgbt community and not just in India but worldwide. The insufficiency of data can be attributed to the fact that lgbtq+ people are marginalized and hence are denied access to jobs, education, health, and other facilities. This in turn implies that this community does not have the required literacy to communicate through social media platforms. It might also be because of an irregular network system across the country. For instance, Hyderabad has a strong and well-built network and support system for the transgender community whereas this is not the case everywhere else. Another reason could be that other platforms are preferred over Reddit in India and are widely used. Thus, it became necessary for us to broaden our scope.

We found a few subreddits that concentrated on the transgender community in India but since the number of users of the subreddits was low, thus the data we scraped was not sufficient enough for us to work it. The first option was to scrape more than one subreddit but the issue with that was with an increase in the number of subreddits, the variety between the data would increase thus leading to varying results. In addition to this, since the number of users in forums dedicated to Indian communities was low, the data would still not be reliable enough to form our analysis. The second option was for us to increase our scope and work not just with the transgender community in India but also with other communities in other parts of the world. This seemed

feasible since we had found the perfect subreddit with appropriate and sufficient data. Therefore, we decided to increase our scope and work with all communities marginalized on the basis of sexuality and gender identities.

The last option was to change our platform from Reddit to Twitter but since the second option worked out pretty well for us we decided to stick with Reddit and continue with our ongoing research.



## Conclusion

This project aimed to understand the lived realities of the marginalized communities during the pandemic through social media data. Reddit was the social media platform used by us to obtain data from the pandemic period. The posts belonged to the period from March, 2020 to August, 2021 and the content of these posts were used by us to perform analysis and reach our conclusions.

These posts were first scraped using Reddit API and Pushshift API and the data obtained was then used to perform LDA topic modeling. Through this modeling, we got the intertopic distance map and relevant terms for each topic. These models were really helpful to perform our analysis and draw conclusions. In addition to this, human analysis was also performed on the data obtained by using Pushshift API.

Our analysis of the entire obtained text was based on the books and the articles read by us to build our foundation on this topic; the same has also been added in the bibliography for further reference.

This project can be a starting point to explore the topic of pandemic experiences in depth. Some additions to this project could have been to explore other social media platforms and compare the results obtained from these different platforms. This addition would have increased the accuracy of the results obtained. Furthermore, a more precise pre-processing of data by manually adding more stopwords could have removed words not relevant to our research but the same has not been done in case our analysis would have benefitted from these terms. As mentioned, this project acts as a starting point and along with the modifications mentioned it could provide us with more accurate results.

It is evident that a few other additions to this project would have provided us with more legitimate conclusions but making additions also meant increasing the scope of the project but due to our time constraints we did not delve much into these modifications.

The current scope of the project also gave us results consistent with our understanding of the marginalized communities on the basis of gender identities and sexualities and we believe the same would happen by extending the scope. Thus, in future we hope to expand the scope and prove our hypothesis correct.

## Bibliography

1. Ahmed, F., & Sifat, R. I. (2021, July 19). Experiences of hijra (transgender) communities during the COVID-19 pandemic in Bangladesh. F1000Research. Retrieved January 11, 2022, from <https://f1000research.com/articles/10-601>
2. Bhalla, R., & Agarwal, S. (2021). Life in a pandemic: Intersectional approach exploring experiences of LGBTQ during COVID-19. *International Journal of Spa and Wellness*, 4(1), 53-68.
3. F, S. (2021, October 18). *Getting to the point with topic modeling: Part 3 - interpreting the visualization*. Alteryx Community. Retrieved May 15, 2022, from <https://community.alteryx.com/t5/Data-Science/Getting-to-the-Point-with-Topic-Modeling-Part-3-Interpreting-the/ba-p/614992#:~:text=The%20intertopic%20distance%20map%20is,each%20topic%20across%20the%20dictionary>.
4. Ganguly, D., & Singh, R. (2021). The Transgender Humanitarian Crisis during the Covid-19 Pandemic in India.
5. Giri. (2022, February 4). *Topic modeling with LDA explained: Applications and how it works*. HDS. Retrieved May 15, 2022, from <https://highdemandskills.com/topic-modeling-intuitive/#:~:text=LDA%20topic%20modeling%20discovers%20topics,based%20on%20a%20Bayesian%20framework>.
6. Kalra, G. (2012). Hijras: The unique transgender culture of India. *International Journal of Culture and Mental Health*, 5(2), 121-126.
7. Kneale, D., & Bécares, L. (2021). Discrimination as a predictor of poor mental health among LGBTQ+ people during the COVID-19 pandemic: cross-sectional analysis of the online Queerantime study. *BMJ open*, 11(6), e049405.
8. Pandya, A., & Redcay, A. (2021). Impact of covid-19 on transgender women and Hijra: Insights from Gujarat, India. *Journal of Human Rights and Social Work*. <https://doi.org/10.1007/s41134-021-00184-y>
9. Philip, B. V. (2021). Impact of covid-19 on transgender persons: The need for an inclusive approach. *International Journal of Sexual Health*, 33(3), 248–267. <https://doi.org/10.1080/19317611.2021.1906375>
10. Roy, R. (2020, July 28). Coronavirus: Kolkata's trans community has been locked out of healthcare and livelihood. Scroll India. <https://scroll.in/article/968182/coronavirus-kolkatas-trans-community-has-been-locked-out-of-healthcare-and-livelihood>
11. Salihefendic, A. (2016, March 19). *How reddit ranking algorithms work*. Medium. Retrieved May 16, 2022, from <https://medium.com/hacking-and-gonzo/how-reddit-ranking-algorithms-work-ef111e33d0d9>
12. Shah, C., Merchant, R., Mahajan, S., & Nevatia, S. (2015). *No outlaws in the gender galaxy*. Zubaan.

13. Stevens, H., & Nikos-Rose, K. (2021, November 5). *LGBTQ+ youth face increased anxiety amid covid-19 pandemic*. University of California. Retrieved May 16, 2022, from <https://www.universityofcalifornia.edu/news/lgbtq-youth-face-increased-anxiety-amid-covid-19-pandemic>
14. Rose, A. (2008). *R/LGBT*. reddit. Retrieved May 16, 2022, from <https://www.reddit.com/r/lgbt/>

## Appendix

The link to the code that has been used by us to create the word clouds and the topic models for this project is as follows :

<https://colab.research.google.com/drive/1ZMfseB29hNhPD75aHUlegedcJGJk8OoZ?usp=sharing>

Comments have been added in the code to provide the users with an understanding of what each snippet of the code performs. The results.csv contains the csv file with all the titles of the post and the models can be seen by running all the cells which can be done by pressing 'Ctrl + F9'.