



Assessing mental health signals among sexual and gender minorities using Twitter data

Yunpeng Zhao^{*}, Yi Guo^{*}, Xing He, Yonghui Wu, Xi Yang and Mattia Prosperi

University of Florida, USA

Yanghua Jin

Zhejiang Gongshang University, China

Jiang Bian

University of Florida, USA

Abstract

Sexual and gender minorities face extreme challenges that breed stigma with alarming consequences damaging their mental health. Nevertheless, sexual and gender minority people and their mental health needs remain little understood. Because of stigma, sexual and gender minorities are often unwilling to self-identify themselves as sexual and gender minorities when asked. However, social media have become popular platforms for health-related researches. We first explored methods to find sexual and gender minorities through their self-identifying tweets, and further classified them into 11 sexual and gender minority subgroups. We then analyzed mental health signals extracted from these sexual and gender minorities' Twitter timelines using a lexicon-based analysis method. We found that (1) sexual and gender minorities expressed more negative feelings, (2) the difference between sexual and gender minority and non-sexual and gender minority people is shrinking after 2015, (3) there are differences among sexual and gender minorities lived in different geographic regions, (4) sexual and gender minorities lived in states with sexual and gender minority-related protection laws and policies expressed more positive emotions, and (5) sexual and gender minorities expressed different levels of mental health signals across different sexual and gender minority subgroups.

Keywords

emotion and mental health, sexual and gender minorities, social media, Twitter

^{*}Equal contribution, co-first authors.

Corresponding authors:

Yanghua Jin, School of Business Administration, Zhejiang Gongshang University, Hangzhou 310018, Zhejiang, China.

Email: jinyanghua@163.com

Introduction

Sexual and gender minorities (SGMs) are minorities of the population whose sexual orientation and gender identity differ from the majority. Over the past decades, a range of civil society organizations have made significant efforts to destigmatize SGM populations, separating being SGM from being mentally ill. However, the stigma persists. For example, the widely adopted International Classification of Diseases (ICD) from the World Health Organization (WHO) continues to consider transsexualism as a diagnosis part of mental, behavioral and neurodevelopmental disorders. On the other hand, studies have consistently documented a high prevalence of mental health distress in the SGM populations. However, mental health issues among SGM individuals remain understudied, mainly due to the lack of population-based representative samples, and hence successful routine health surveillance efforts. One key issue that limits the availability of representative samples, as identified in a recent *Lancet* review paper,¹ is the “Lack of standardized survey items on population-based surveys to identify transgender respondents.” Nevertheless, we think the issue lies beyond just standardizing survey items on population-based surveys. SGMs face systematic social and economic marginalization, pathologization, stigma, discrimination, and even violence; they are often unwilling to self-identify as SGMs when asked and reluctant to participate in traditional surveys.

Meanwhile, social media have become increasingly popular communication platforms for health-related studies and brought rapid changes to the health communication landscape. Twitter, in particular, has been used to promote healthy behavior,² improve medical and patient education,³ and recruit research participants,⁴ including from vulnerable populations such as the SGMs.^{5,6} Furthermore, individuals are also voluntarily sharing a critical amount of personal information, including their health experience in various online communities such as Twitter, blogs, and topic-specific discussion forums. These user-generated social media data provide unique insights into public and consumer health and are invaluable data sources for understanding various social and health issues. Twitter, as a good example, can serve as a data source for analyzing public health-related problems such as influenza infection,⁷ surveillance for cardiac arrest,⁸ tobacco use,⁹ prescription drug abuse,¹⁰ and measles outbreak.¹¹

Nevertheless, despite that social media platforms such as Twitter have been widely used as a communication tool to deliver health intervention programs targeting the SGM populations,^{5,6} there have been very few studies that used these rich user-generated health data to understand SGMs' health statuses and health behavior. In fact, we performed a literature review and revealed that there is no published study on the discussion of mental health issues in the SGM populations using social media data. To fill this gap, we have conducted two studies on the self-identification and mental health problems of the SGM populations based on Twitter data. In the first study, we developed a machine learning technique (i.e. random forest (RF)) to distinguish self-identified SGMs Twitter users from general Twitter users via analyzing their tweets and assessed the variance of the gender identification terms these SGM Twitter users used to describe themselves.¹² In a follow-up study, we explored emotional and mental health signals of these SGMs using the same dataset.¹³

In this study, we first collected significantly more Twitter data relevant to SGM extending our original dataset to a total number of 35,053,757 tweets. Further, we extracted the geolocation information of these tweets and the corresponding users, which allowed us to conduct a more comprehensive analysis of the mental health signals (i.e. related affect processes such as positive or negative emotions, anger, anxiety, and sadness) exposed through SGM Twitter users' tweets. More specific, we aim to answer the following five research questions (RQs) in this study, while in our previous study,¹³ we only addressed two research questions (RQ1 and RQ5) with a significantly smaller dataset.

RQ1. Do SGM individuals experience different affect processes (i.e. emotional states, including positive and negative emotions, anxiety, anger, and sadness) compared with non-SGM people when discussing gender identity and sexual orientation issues?

RQ2. Do SGM individuals' affect processes change by time when discussing gender identity and sexual orientation issues?

RQ3. Do SGM individuals experience different affect processes across different geographic locations when discussing gender identity and sexual orientation issues?

RQ4. Do SGM-related protection laws and policies have any impact on SGM individuals' affect processes?

RQ5. Do different SGM subpopulations (e.g. gay vs lesbian and transman vs transwoman) experience different affect processes when discussing their gender identities and sexual orientations?

Method

The general idea underlying our approach is to collect tweets that are relevant to the discussion of SGM-related issues, determine self-identifying tweets and Twitter users (i.e. users who have stated that they were SGMs), and assess these users' emotional states using an established text analysis approach. Figure 1 depicts an overview of our 5-step analysis workflow. In the following sections, we describe each of the steps in detail.

Step 1: data collection

The data used in this study are from four different sources: (1) the first dataset is from our 2015 study¹² and the tweets were crawled with *tweetf0rm*,^{14,15} a Python tool we developed, which can access the Twitter Application Program Interface (API) and collect targeted Twitter data based on a list of keywords. The list of keywords was mainly related to gender identification terms such as “transwomen” and “genderqueer” as well as a set of keywords that indicated relevance of the tweets to SGM discussions such as “testosterone” and “gender reassignment surgery.” The list of keywords was developed through a snowball sampling process, where we started with a list of obvious seed keywords collected from online information sites relevant to definitions of SGM identity terms (e.g. <http://transstudent.org/definitions>). We then iteratively queried (a) our collection of historical public random tweets (i.e. collected through Twitter streaming API) and (b) Twitter's own web search interface (i.e. <https://twitter.com/search-home?lang=en>), and manually reviewed randomly selected sample tweets to discover new SGM-related keywords (i.e. that co-occur with one of the existing keywords, but were not in the existing keyword list) until no new keywords were found. Through this process, we found 344 keywords in total in our original 2015 study, (2) from 28 November 2017 to 6 June 2018, we used a subset of these keywords (i.e. 27 most commonly used SGM-related keywords) to collect more SGM-related tweets using *tweetf0rm*, (3) we extended our 2015 keyword list to a total of 640 keywords through the same snowball sampling process in order to adapt to the changes in people's vocabulary on the Internet. We used this extended keyword list to collect more Twitter data from 6 May 2018 to 7 July 2018, and (4) we also used this extended keyword list to search for SGM-related tweets on a database of public random tweets, which was collected using the Twitter streaming API with tweets from July 2006 to December 2017 (i.e. with a few months of gaps).

Step 2: data preprocessing

We preprocessed the collected data to eliminate tweets that (1) were non-English or (2) not posted in the United States (i.e. users who cannot be geotagged to a US state). The geolocation of the Twitter user was extracted (i.e. with a resolution to US states) by using a Twitter

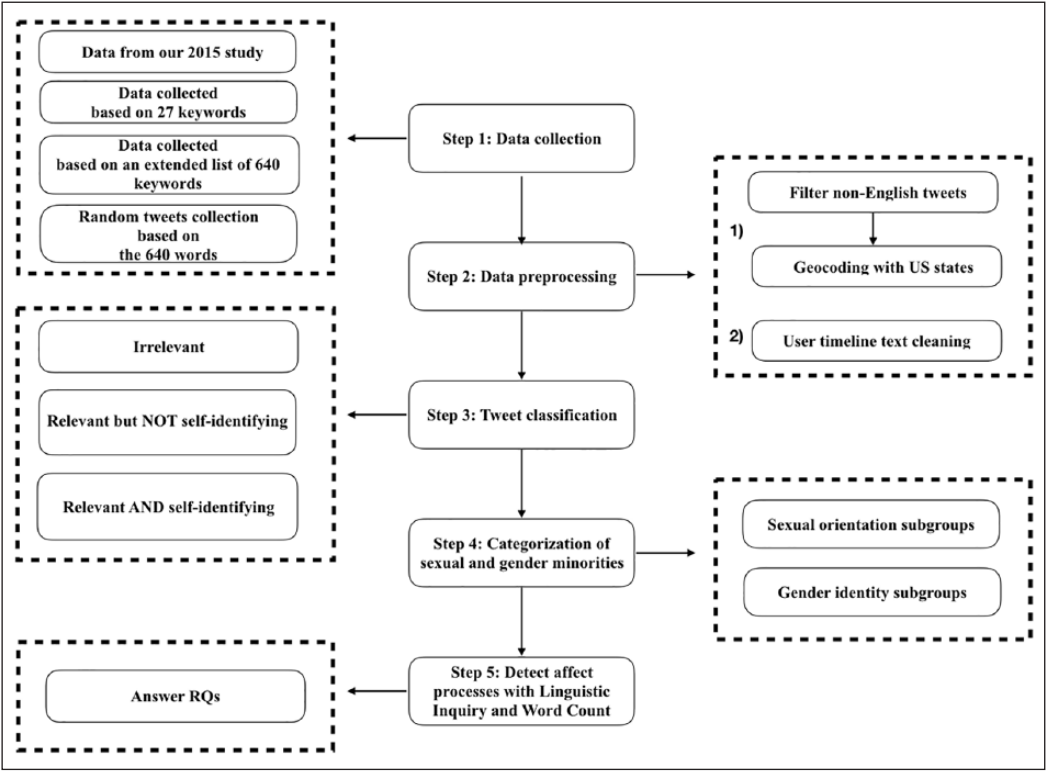


Figure 1. The general analysis workflow.

geocoder tool we developed.¹² In Twitter, there are several ways that geographic information can be attached to a tweet: (1) if a GPS-enabled (i.e. GPS, Global Positioning System) in a mobile device, a pair of latitude and longitude will appear in a tweet; (2) the associated user profile can be geocoded (either to a specific GPS location or a geographic ‘place’; and (3) the user might have filled the geographic information manually in the “location” attribute with free-text. Specifically, if the geocodes (i.e. GPS coordinates or a “place” tag) were available, we attempted to reverse geocode to a state name; otherwise, the Twitter geocoder tool will match the “location” string with a number of lexical patterns such as the name of a state (e.g. Texas or Florida), or a city name in combination with a state name or state abbreviation in various possible formats (e.g. “—, fl” or “—, florida” or “—, fl, usa”). We discarded the tweets that cannot be assigned to a specific US state. We also made a number of other efforts to clean up the tweets for further text analysis steps: (1) removed hyperlinks (e.g. “https://t.co/xyz”), (2) removed mentions (e.g. “@username”), and (3) converted hashtags into original English words (e.g. converted “#gay” to “gay”).

Step 3: tweet classification

Even though a tweet contains one or more of the SGM-related search keywords, the tweet may not be relevant to the SGM discussion (e.g. “trans” could mean either “transgender” or “transmission”). Thus, we developed a two-step process and built two classification models to categorize the

massive tweets that we collected into 3 groups (i.e. irrelevant, relevant but NOT self-identifying, and relevant AND self-identifying).

We reused the annotated data (i.e. 6,058 tweets manually labeled by two raters) from our previous study¹² as the training set to build these classifiers. The training dataset were manually annotated by three people and each tweet was classified into one of three labels: “irrelevant” (661 tweets), “relevant but NOT self-identifying” (4,619 tweets), and “relevant AND self-identifying” (778 tweets). When disagreements between the three annotators occurred, we used the majority rule to determine the final label. The inter-rater agreement (i.e. Fleiss’ kappa) was 0.96. Different from our previous study, we experimented with three different machine learning algorithms namely RF, support vector machines (SVM), and convolutional neural networks (CNNs). We implemented the RF and SVM classifiers via the scikit-learn library¹⁶ and used the Term Frequency–Inverse Document Frequency (TF-IDF) scheme to convert each tweet into a feature vector. Stop words were removed from the tweets. We optimized the hyper-parameters of the RF and SVM models using 10-fold cross-validation. We implemented the CNN-based deep learning model in Keras¹⁷ on top of the Tensorflow¹⁸ framework. The architecture of the CNN model included an embedding layer, a convolutional layer, a global max pooling layer and a sigmoid output dense layer. We initialized the embedding layer with the GloVe pretrained 200 dimension Twitter word embeddings.¹⁹ In the convolutional layer, we set the number of filters to 64, the length of filter to 3, and the dropout rate to 0.2. We experimented with different settings—the dimension (i.e. 50, 100, 200) of the pretrained Twitter word embeddings, the number (i.e. 16, 32, 64) of the filters, and the length (i.e. 3, 5, 7) of the filters, and the dropout rate (i.e. from 0.2 to 0.7 with 0.1 increment)—to find the best-performing hyper-parameters. To prevent model overfitting, we applied an early stop strategy based on the value loss on the validation set.

Step 4: categorization of SGMs

We integrated the four different data sources and eliminated duplicates across them. We then used the two classification models built in Step 3 to find all self-identifying tweets, and subsequently their self-reported gender identities and/or sexual orientations. Furthermore, two annotators manually validated these self-identifying tweets and categorized these SGMs into their corresponding SGM subgroups as shown in Figure 2. The inter-rate agreement between the two annotators was 0.93. Disagreements were resolved by a third annotator.

Step 5: detect affect processes with linguistic inquiry and word count

The linguistic inquiry and word count (LIWC) is a validated text analysis tool that can exam the language people used in a given text and reveal their thoughts, feelings, personality, and motivations. With a given text, it counts the percentage of words that reflect different emotions, thinking styles, social concerns, and sentiments of the writer. In a recent systematic review, Wongkoblap et al.²⁰ reported that the LIWC is the most popular tool used to extract potential signals of mental problems from social media data. There are totally 86-word categories in LIWC, where 41 out of them are psychological constructs (e.g. affect processes, cognition, biological processes, and drives). The affect process constructs (i.e. positive emotion, negative emotion, anger, anxiety, and sadness) are closely related to an individual’s mental health.

To answer the five RQs, we applied the LIWC tool on the entire Twitter timelines (i.e. the complete tweet history of each Twitter user of interest) of the corresponding sexual orientation and SGM gender identity groups and compared the affect processes of the different user groups as reflected in their tweets.

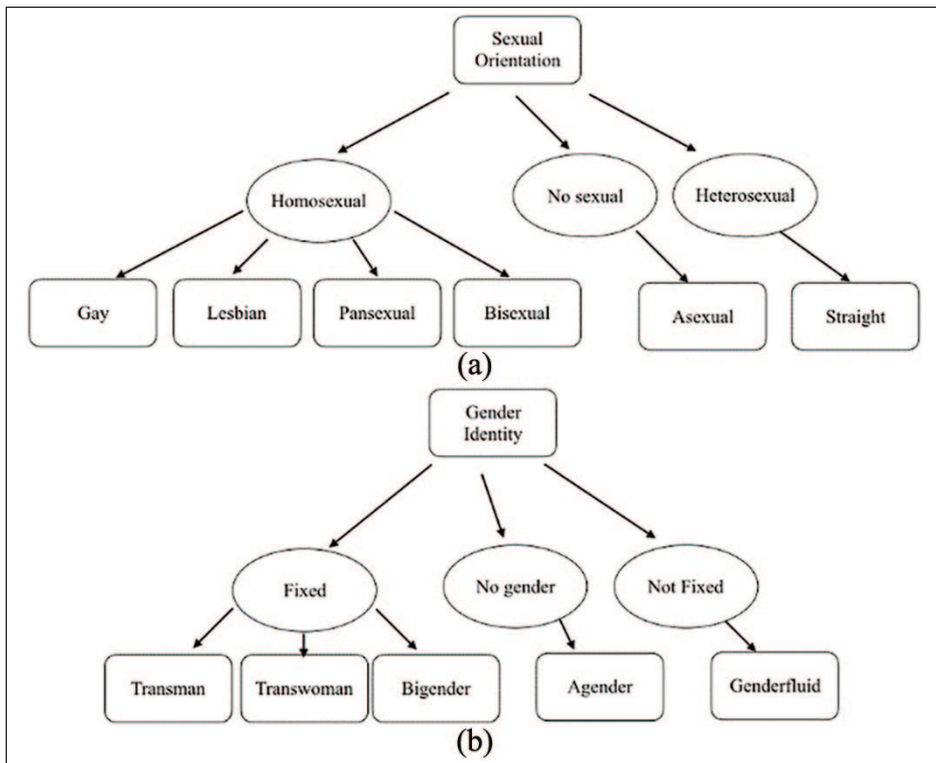


Figure 2. The categorization of SGMs' sexual orientations and gender identities, where (a) shows the categorization of sexual orientations (i.e. gay, lesbian, bisexual, pansexual, asexual, and straight) and (b) shows the different gender identity categories (i.e. transman, transwoman, bigender, genderfluid, and agender).

Results

Data sources

In this study, our data came from four different sources as shown in Table 1. First, in our previous study,¹² we collected 56,653,010 tweets using 344 keywords from 17 January 2015 to 12 May 2015. After filtering out non-English tweets, there were 11,502,289 tweets left, which we retained for further analysis. Second, we collected 5,607,182 tweets from 28 November 2017 to 6 June 2018 using a set of 27 commonly used SGM-related keywords. After filtering out non-English tweets, there were 3,889,191 tweets left. Third, we extended the original list of 344 keywords from our previous study to 640 keywords to adapt to the changing terms used on the Internet related to SGM. We collected our third dataset using this extended 640 keyword list from 6 May 2018 to 22 June 2018. We collected 25,823,163 tweets in total. After filtering out non-English tweets, there were 9,974,721 tweets left. Fourth, we also used the extended 640 keywords to search on a database with random public tweets from July 2006 to December 2017, which we collected using the Twitter steaming API. We found 3,623,412 SGM-related tweets from this random tweet database, within which 2,544,593 tweets were written in English.

After integrating and eliminating duplicates across the four different data sources, there were 27,303,446 unique tweets from 7,199,474 unique users, in which 2,296,923 tweets were in English and 895,723 users' geo locations can be identified.

Table 1. The Twitter data from four different sources.

Data source	Data time range	Number of keywords ^a	Number of tweets before preprocessing	Number of English tweets
Data from our 2015 study	10/01/2015 to 12/05/2015	344 keywords	56,653,010	11,502,289
Data collected based on 27 keywords	28/11/2017 to 08/06/2018	27 most common keywords related to SGM	5,607,182	3,889,191
Data collected based on an extended list of 640 keywords	06/05/2018 to 22/06/2018	640 keywords	25,823,163	9,974,721
Random tweets collection based on the 640 words	07/2006 to 12/2017	640 keywords	3,623,412	2,544,593

^aAll the keywords are listed in Supplemental Appendix A. Part 1 lists the 344 keywords, part 2 lists the 640 extended keywords, and part 3 is the 27 keywords.

The three classification models for identifying relevant SGM tweets and self-identified SGM Twitter users

We explored three methods (i.e. random forests, support vector machines and CNNs) to build supervised models for classifying the tweets. In general, the random forests and support vector machines have good out-of-box performance for most classification tasks, including on text data,^{21,22} while recent reports have shown state-of-the-art performance with deep learning models such as the CNNs surpassing traditional machine learning methods.²³ Table 2 shows the performance of the different classifiers and classification tasks. We used 80% of the annotated data (i.e. 6058 tweets manually labeled by two raters) for training and the performance metrics were measured on the rest 20% as the independent test data. As shown in Table 2, the random forest models outperformed the CNNs and support vector machine models for both tasks. Thus, we adopted the random forest models as the final classifiers.

Since we reused the annotated data from our previous study¹² to build these classifiers, we also validated the classifiers' performance on new data. We randomly selected and annotated 150 tweets from our new dataset and manually reviewed each tweet to create the gold-standard labels. The model performance on the first classification task (i.e. relevant vs irrelevant) has a precision of 0.95, a recall of 0.95, and a F-measure of 0.95. The performance on the second classification task (i.e. self-identifying vs NOT self-identifying) has a precision of 0.89, a recall of 0.86, and a F-measure of 0.86.

Categorization of SGMs Twitter users

The best performed classifiers (i.e. the random forest model) identified 12,540 self-identifying tweets from 8,439 unique Twitter users. We manually evaluated the self-identifying tweets from these 8,439 users, and found that 7,033 (1,406 false positives) were correctly identified by the classifiers. Within these self-identifying Twitter users, 5,066 users claimed themselves having non-binary genders (i.e. beyond traditional male or female biological sex) and 3,180 users claimed they were sexual minorities. Among these users, 1,213 users discussed both their gender identities and sexual orientations in their Twitter posts. We then further classified these SGMs who self-reported their specific gender identities and sexual orientations into 11 subgroups (i.e. gender identities:

Table 2. The performance comparison of the classifiers.

Classifier	Precision	Recall	F-score
Task 1: relevant vs irrelevant			
CNN	0.97	0.97	0.97
Random forest	0.98	0.98	0.98
Support vector machines	0.97	0.97	0.97
Task 2: self-identifying vs NOT self-identifying			
CNN	0.84	0.85	0.84
Random forest	0.88	0.89	0.86
Support vector machines	0.87	0.88	0.86

CNN: convolutional neural network.

transman, transwoman, bigender; genderfluid, and agender; and sexual orientations: gay, lesbian, bisexual, pansexual, asexual and straight). Two researchers annotated the self-report SGM users. The Cohen kappa was 0.93. Conflicts were resolved by third person. These subgroups and the number of people in each group are shown in Figure 3. We also show example tweets of each group in Table 3. Note that the goal of our previous study¹² was to find self-identifying tweets (i.e. a tweet where the user explicitly expressed their sexual orientations including those who are straight). Thus, our dataset included users who were self-identified as straight (125 users). These users were not considered as SGMs and excluded from the LIWC analyses for RQ1-4. However, these users were included in the analysis for RQ5 as a comparison group of different sexual orientations.

Mental health signals of the self-identified SGMs on Twitter

To reliably assess individual users’ mental health signals (i.e. affect processes) using LIWC, we first collected SGM population groups’ Twitter timeline (i.e. so that we have enough data samples on individual users). We then applied the LIWC tool on the different SGM population groups’ Twitter timelines to answer the five RQs. Specially, we used LIWC to analyze the affect processes related to mental health signals such as positive and negative emotions (also called sentiments), anger, anxiety, and sadness expressed in different SGM population group users’ tweets.

To answer RQ1, we attempted to collect the 7,033 users’ timelines; however, only 6,604 out of the 7,033 users’ timelines can be retrieved (e.g. non-public users’ timelines cannot be collected, since it needs authorization from the users). After excluding those who were self-identified themselves as straight (125 users), the SGM case group has 6,489 users in total. We then randomly selected a control group with 6,489 users matched on corresponding geographic locations (i.e. with a resolution to the US states) from the “Relevant but NOT self-identifying” user group and collected their Twitter user timelines as well. Table 4 shows the basic statistics of the collected user timelines.

We treated each user timeline as a single document and fed all user documents into the LIWC tool. The LIWC tool then calculated a score (percentage of words) for each affect process (i.e. emotion states, anxiety, anger, and sadness). These scores reflect the severity of the measures (e.g. a higher positive emotion score indicates that the user has exhibited more positive thoughts in her tweets). Figure 4 shows the comparison of these different emotional states between SGM vs non-SGM groups. To assess whether our results were significant, we also performed Student’s t-tests of the LIWC scores between SGM and non-SGM groups for each affect process category. As shown in Table 5, except for anxiety, there were significant differences between the SGM group and the

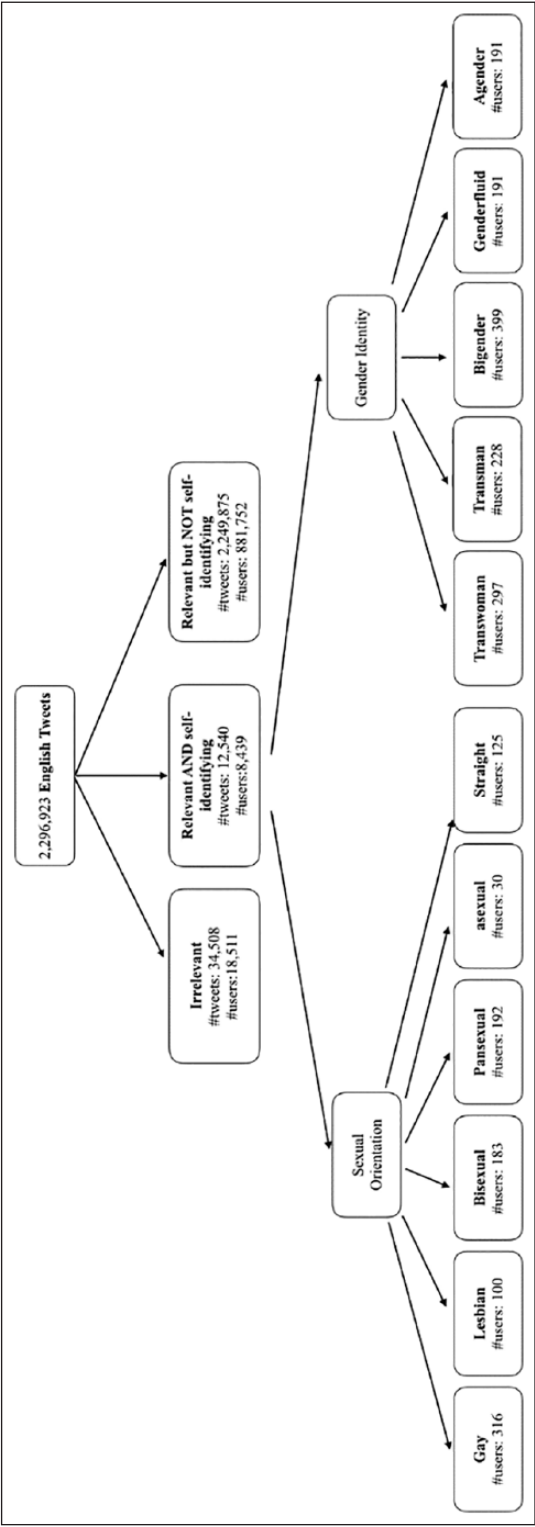


Figure 3. Results of categorizing the Twitter users into sexual orientation and gender identity subgroups.

Table 3. Examples of self-identifying tweets.

Group	Example tweets
Gender identity subgroups	
Transman	<ul style="list-style-type: none"> • "I'm not enough of a guy for gay guys, I tell myself even tho I identify as a trans guy atm Am I fucked up or what" • "I felt so ashamed that it led me to attempt suicide. After waking up in the hospital the next day, I decided to start living my truth. Since then, I have come out as a transman."
Transwoman	<ul style="list-style-type: none"> • "im trans women to me like i'm one of your bros" • "I am a Trans-Woman, I am proud of it. I hope Trump sees this, and leave us a lone."
Bigender	<ul style="list-style-type: none"> • "I recognize myself as a pansexual bigender" • "@LauraJaneGrace I just came out as bigender to my mom...didn't go super well. But I'm just happy to know I can find acceptance through AM!"
Genderfluid	<ul style="list-style-type: none"> • "I'm pansexual, genderfluid and trans and I'm proud to be who I am. #PRIDE #PrideMonth" • "And when I explained my gender fluidity to my husband he understood and supported me and that was super amazing and great!"
Agender	<ul style="list-style-type: none"> • "I am agender, not a gender" • "i love how most of my gender headcanons tend to lean towards agender/nonbinary ... bc ... that me tbh."
Other ^a	<ul style="list-style-type: none"> • "Today's the one year anniversary of the day I came out to my mom as transgender" • "yes im transgender. am i ashamed of it? no not at all. no, i am not a "dyke" or "lesbian" i respond to male bc that's how i see myself."
Sexual orientation subgroups	
Gay	<ul style="list-style-type: none"> • "I am gay, I am trans, and I am tired." • "I'm learning so much! I'm gay myself, but I nvr rly understood much about the trans lifestyle. This is eye-opening. <url>"
Lesbian	<ul style="list-style-type: none"> • "im so tired of arguing abt identity politics lol. i shouldnt have to justify myself to others. im nonbinary lesbian End Of" • "I now identify as a lesbian and yes just because I had a bad experience it doesn't mean I hate men. I'm just not attracted to them. Sexuality is very much fluid. Maybe one day I'll suddenly be pansexual. Idk. All I know is I love women, I love myself, this is me. #PrideMonth"
Bisexual	<ul style="list-style-type: none"> • "I've always been bi-curious but it's time for me to be completely honest wih myself. I am bisexual and proud. #LGBTQ <url>" • "I call myself queer/bisexual and I define it as being attracted both to people of my gender and to people not of my gender. None of that implies that there are only 2 genders ..."
Pansexual	<ul style="list-style-type: none"> • "I'm breadlad because I'm PANsexual get it ha ha ha ha happy pride month.#PrideMonth" • "I love my pansexual fave"
Asexual	<ul style="list-style-type: none"> • "I'm asexual. I dare myself bc I'm so charming <url>" • "@username I classify myself as asexual"
Straight	<ul style="list-style-type: none"> • "also, i am a cis woman, so i know i'm speaking from my own limited perspective. the trans experience is not my own. but i have to wonder ..." • "@username I'm cis and I hate myself"
Other ^a	<ul style="list-style-type: none"> • "Im a homosexual sir <url>" • "im worried spotify knows im a homosexual <url>"

SGM: sexual and gender minority.

^aSelf-identify tweets that belong to none of the SGM subgroups.

Table 4. The numbers of users and tweets of SGM versus non-SGM groups.

	Number of users	Number of tweets, mean (min—max; std)	Number of days, mean (min—max; std)
SGM	6489	3032 (1–9624; 1835.60)	781 (1–3,987; 810.14)
non-SGM	6489	2421 (1–6555, 1132.12)	742 (1–3870, 823.32)

SGM: sexual and gender minority.

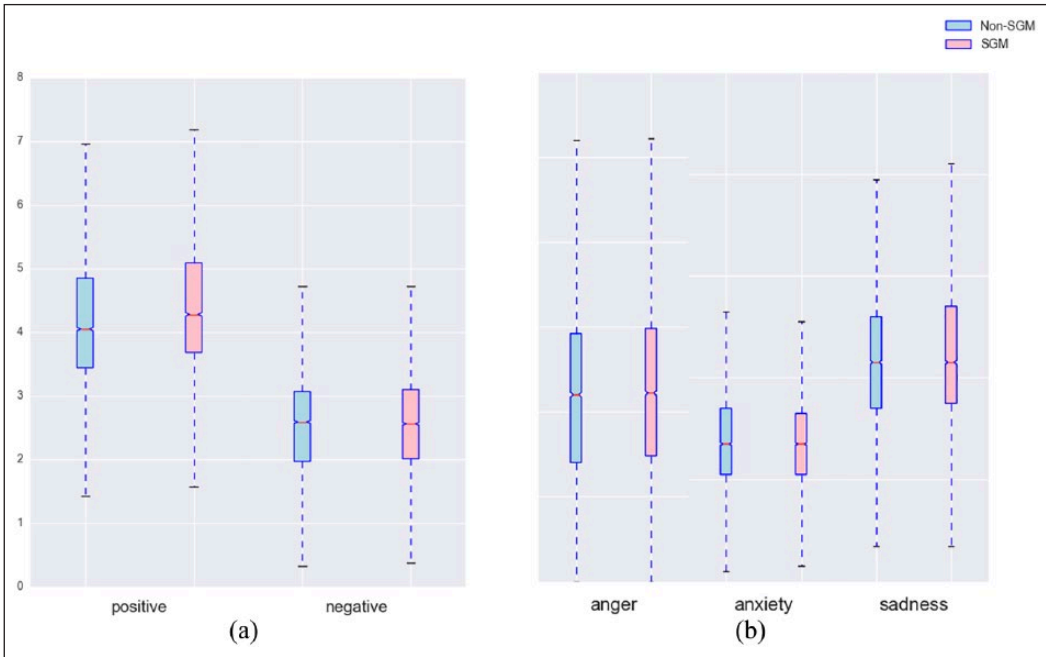


Figure 4. A comparison of SGM vs non-SGM Twitter users' affect processes expressed in their tweets ((a): positive vs negative and (b): anger vs anxiety vs sadness).

non-SGM group in terms of these affect processes. In short, the SGM group not only has both higher negative and positive emotion scores, but also has expressed more anger, more anxiety, and more sadness issues in their tweets.

To answer RQ2, we sliced SGM users' timelines by year and random selected the equal number of people from the "Relevant but NOT self-identifying" user group from each corresponding year. Table 6 shows the statistics of the case and control groups. We selected the equal numbers of users from the non-SGM control group corresponding to the number of users in the SGM case group for each year. Figure 5 shows the LIWC score comparison between the SGM group and non-SGM group by year. As shown in Figure 5, for negative emotion, sadness, and anger, the LIWC score differences between SGM and non-SGM groups are shrinking after 2015.

To answer RQ3, we compared the LIWC scores of the affect processes between SGM and non-SGM groups by state. There are 8 states (i.e. Hawaii, Illinois, Massachusetts, Michigan, Missouri, North Dakota, Tennessee, and Washington) where SGMS used words in at least one or more than one of negative emotions (i.e. negative emotion, anger, anxiety, and sadness) than

Table 5. The t-test results of comparing the five different mental health signals between SGM and non-SGM groups.

	Positive emotion, mean (min—max; std)	Negative emotion, mean (min—max; std)	Anger, mean (min—max; std)	Anxiety, mean (min—max; std)	Sadness, mean (min—max; std)
SGM	4.54 (0.0–45.53; 1.53)	2.56 (0.0–12.12; 0.89)	1.15 (0.0–10.61; 0.62)	0.28 (0.0–3.05; 0.14)	0.45 (0.0–7.14; 0.23)
non-SGM	4.29 (0.0–31.30; 1.45)	2.51 (0.0–12.32; 0.92)	1.14 (0.0–12.12; 0.60)	0.28 (0.0–3.05; 0.14)	0.44 (0.0–4.17; 0.20)
p-value	<0.001	<0.001	<0.001	0.96	<0.001

SGM: sexual and gender minority.

Table 6. The numbers of users and tweets of SGM versus non-SGM groups by year.

Year	Number of users in SGM case and non-SGM control groups	Number of tweets by SGMs, mean (min—max; std)	Number of tweets by non-SGMs, mean (min—max; std)
2012	673	332 (1–2864; 435.59)	207 (1–2658; 310.00)
2013	1025	418 (1–3100; 494.13)	231 (1–2522; 314.97)
2014	1769	714 (1–3094; 768.60)	278 (1–2607; 344.43)
2015	2843	982 (1–9394; 1392.21)	354 (1–3198; 424.79)
2016	2479	501 (1–3,203; 560.18)	483 (1–3242; 512.63)
2017	4134	782 (1–3210; 688.35)	833 (1–3221; 694.38)
2018	5849	1399 (1–9363; 1282.46)	1233 (1–4565; 1,126.60)

SGM: sexual and gender minority.

non-SGMs ($p < 0.05$). The list of US states which had significant p value of each affect process was shown in Table 7; while we did not observe such effect in other states (i.e. possibly due to the small sample sizes of the tweets collected for other states). An example is shown in Figure 6 for the Washington state.

We also compared mental health signals of the SGM groups across states. We created US state heatmaps of each mental health signal as shown in Figure 7. The intensity of the color is positively proportional to the score. SGM people in New Mexico expressed more positive emotions in their tweets than any other states. SGM people in Alaska had the highest score of negative emotion and anger. SGM people in Arkansas had the highest score of anxiety. Although there were no big differences of the sadness scores across states, SGM people in Rhode Island had the highest score of sadness.

To answer RQ4, we considered whether the existence of SGM-specific protection laws and policies (i.e. education discrimination law, healthcare discrimination law, housing discrimination law, employment discrimination law, hate crimes law, public accommodations law, conversion therapy and gender on id policy)²⁴ have any impacts on the affect processes of the SGM Twitter users. For each policy, we grouped SGM users living in states with the policy as the case group; while the SGM users living in states without the policy as the control group. We then compared the LIWC scores of the affect processes between the case and control groups. For example, Figure 8 shows the comparison between SGM users living in the states with and without education discrimination laws; however, only negative emotion and sadness scores were statistically significant ($p < 0.05$) as shown in Table 8.

SGM people expressed more anger in the states without employment, housing, conversion therapy, and public accommodations laws. There was no difference in any affect processes between SGM people living in states with and without a gender on id policy.

To answer RQ5, we further evaluated the LIWC affect process scores by the different sexual orientation and gender identity subgroups. As shown in Figure 9, users with different gender identities (i.e. transman, transwoman, bigender, genderfluid, and agender) clearly expressed different levels of affect processes in their Twitter timelines. For example, agender people used less positive emotion, more negative emotion, anxiety, and anger words. Figure 10 shows the LIWC affect process scores across users with different sexual orientations (i.e. gay, lesbian, bisexual, pansexual, asexual, and straight). Bisexual individuals expressed more positive emotions; pansexual individuals used more negative emotion, anger, and sadness words; while straight individuals used more words associated with negative emotion, anxiety, and anger.

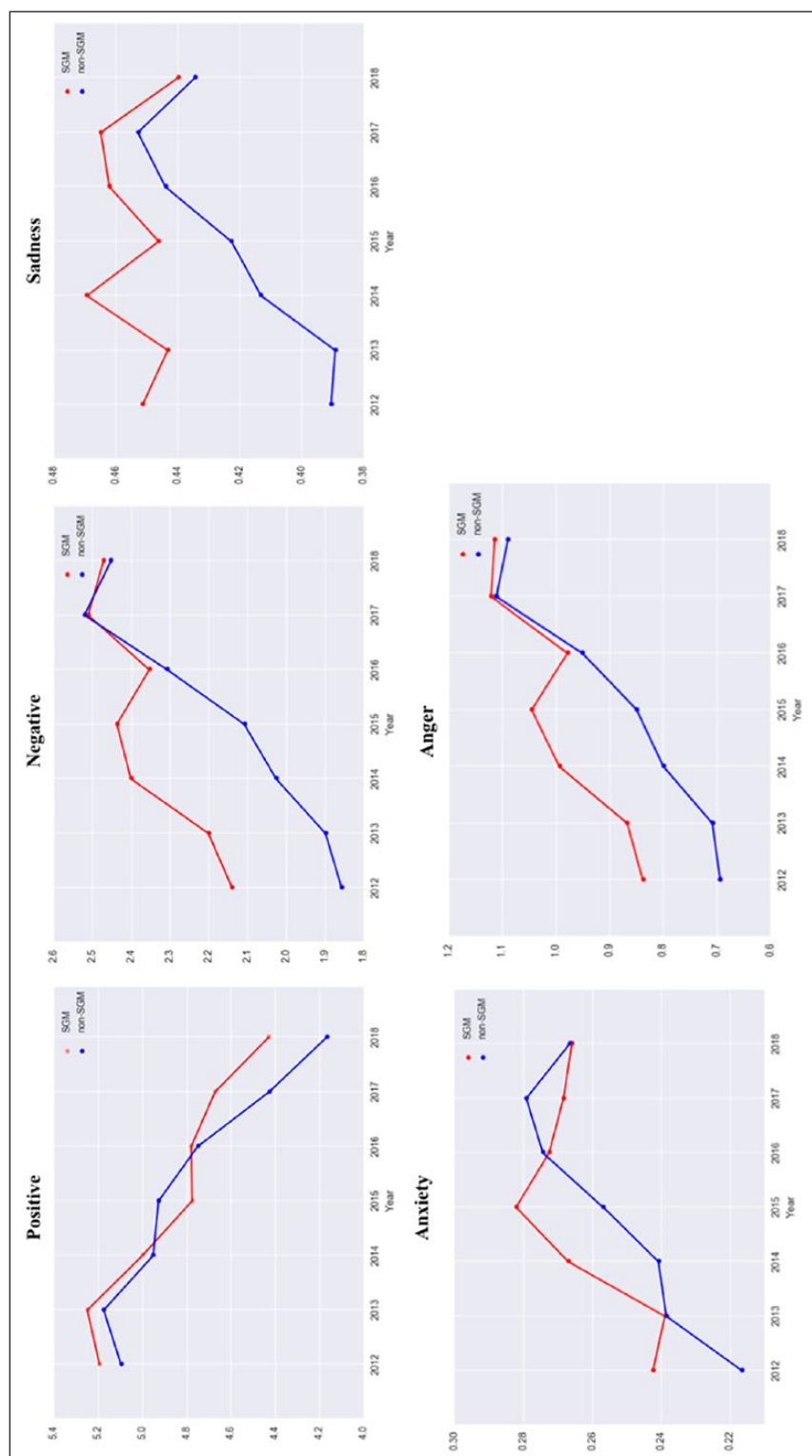


Figure 5. Affect processes comparison between SGM group and non-SGM group by year.

Table 7. The list of US states which had significant p-value of each affect process.

Affect process	The name of US states
Positive	Florida, California, Maryland, New York, Ohio, Washington, D.C., Virginia, Pennsylvania, Missouri, Wisconsin, Washington, Rhode Island
Negative	Michigan, Washington, Missouri, Idaho ^a
Anxiety	Washington, Tennessee
Anger	Michigan, Illinois, North Dakota, Missouri, Massachusetts
Sadness	Hawaii

SGM: sexual and gender minority.

^aIdaho is the only state where SGM population expressed less negative emotion than non-SGM group. All other states are the places where SGM population expressed more negative emotion than non-SGM group.

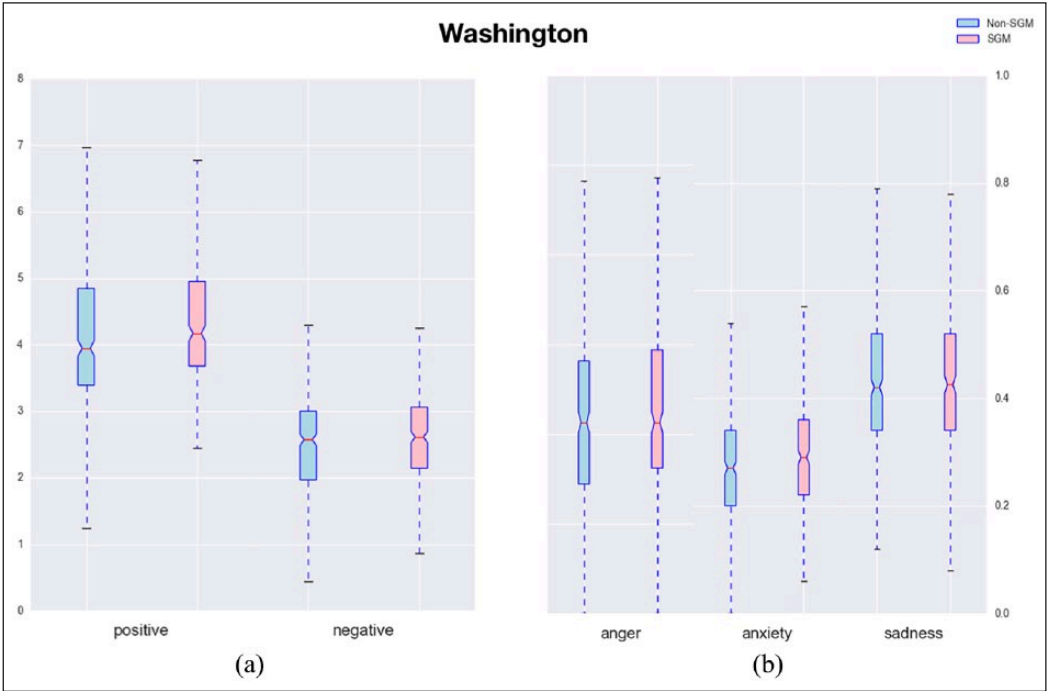


Figure 6. A comparison of SGM vs non-SGM Twitter users' affect processes expressed in their tweets in Washington state ((a): positive vs negative and (b): anger vs anxiety vs sadness).

Discussion

SGM people face extreme challenges (e.g. discrimination, human right issues) from the societies. Stigma as well as prejudice edged them toward the margins of societies, leading to discrimination and abuse, with alarming consequences damaging not only their physical but more significantly their mental health. A news article from the Center for American Progress²⁵ in 2017 showed 68.5 percent of SGM people reported that discrimination at least somewhat negatively affected their psychological well-being and 47.7 percent of SGM people reported that discrimination negatively impacted their spiritual well-being. Nevertheless, SGM people and their health needs, especially their mental health

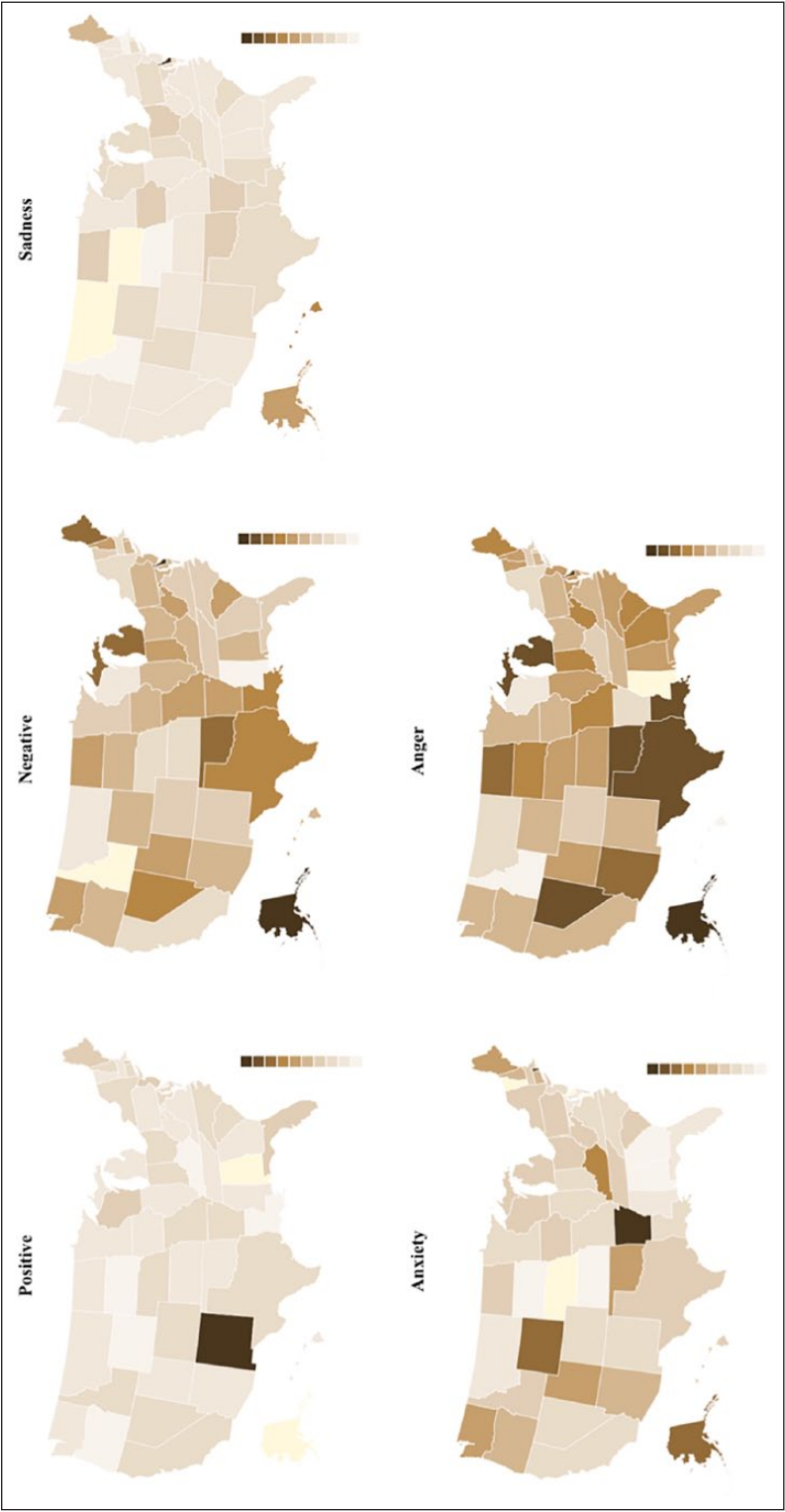


Figure 7. Comparison of the affect processes of SGM Twitter users across states.

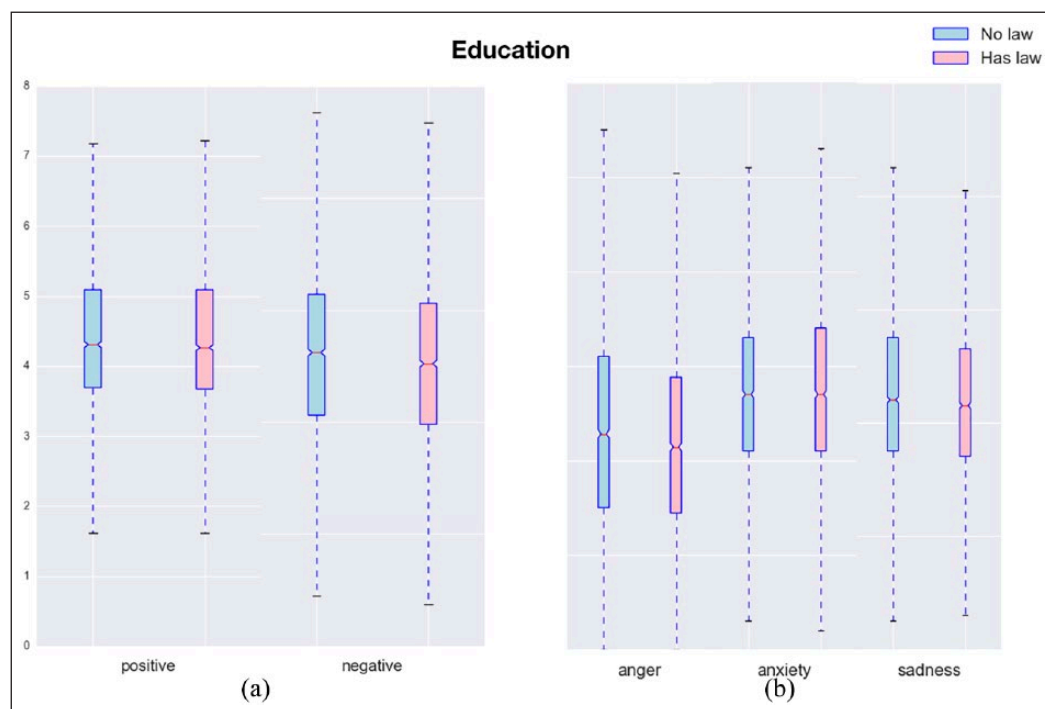


Figure 8. The affect processes ((a): positive vs negative and (b): anger vs anxiety vs sadness) between SGM individuals living in states with versus without anti-education discrimination law.

needs, remain little understood, not only by health-care providers but also more generally in society. Few population-level data exist with which to monitor the health of SGM people in the United States, because routine national health surveillance does not assess gender identity as an equity startifier.¹ Although there exists the US Transgender Survey (USTS, previously known as the National Transgender Discrimination Survey) that aimed to quantify the discrimination and violence SGM people face, these surveys might not capture the full spectrum of SGM people's health challenges. In particular, because of stigma and discrimination, SGM people are often unwilling to self-identify as SGMs when asked and reluctant to participate in these traditional surveys, especially on sensitive topics, such as mental health issues. Social media data, on the other hand, provide a golden opportunity to help us understand vulnerable populations such as SGMs.

Built upon the data we previously collected, we extended the dataset by integrating four different data sources and developed machine learning models to identify SGM individuals through finding their self-identifying tweets. We then applied a lexicon-based method (i.e. the LIWC tool) to extract affect processes that indicate individuals' mental health statuses (i.e. positive and negative emotions, anger, anxiety, and sadness) from different SGM subpopulations' Twitter timelines to answer five specific research questions. Our results suggested that SGM individuals expressed more negative feelings (i.e. negative emotion, anxiety, anger and sadness) in their tweets compared with non-SGM people. This fact also reflected from a survey²⁶ from the Pew Research Center in 2013 that SGM population and the general public are notably different in the happiness self-evaluation. Only 18% of SGM people described themselves as "very happy" when they were asked to evaluate their happiness, while the percentage was 30% of people in the general public. Negative

Table 8. The t-test results of SGM users living in states with versus without SGM specific anti-education discrimination laws.

	Positive emotion, mean (min—max; std)	Negative emotion, mean (min—max; std)	Anger, mean (min— max; std)	Anxiety, mean (min— max; std)	Sadness, mean (min—max; std)
SGM group living in states with anti-education discrimination law	4.54 (0.0–45.53; 1.66)	2.51 (0.0–10.71; 0.86)	1.11 (0.0–2.22; 0.55)	0.27 (0.0–4.91; 0.12)	0.44 (0.0–6.25; 0.23)
SGM group living in states without anti-education discrimination law	4.55 (0.0–21.47; 1.37)	2.61 (0.0–12.12; 0.90)	1.19 (0.0–3.05; 0.65)	0.28 (0.0–12.12; 0.14)	0.46 (0.0–7.14; 0.24)
p-value	0.88	<0.001	0.96	0.32	0.004

SGM: sexual and gender minority.

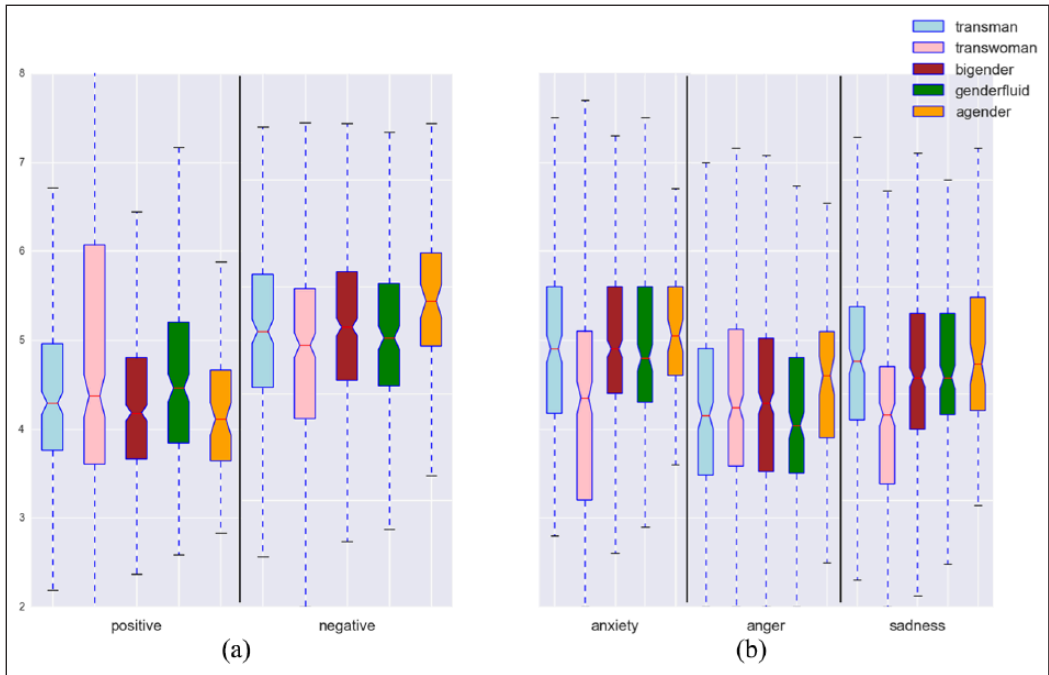


Figure 9. The affect processes ((a): positive vs negative and (b): anger vs anxiety vs sadness) expressed in SGM individuals' tweets across different sexual orientation groups.

affect processes can dampen their enthusiasm for life and longtime holding of negative emotions can cause a downward spiral.

Furthermore, we found the LIWC score differences of negative emotion and sadness between the SGM group and non-SGM group were shrinking after 2015. This phenomenon might be attributed to the national and global efforts in improving human right of SGMs. For example, legally sanctioned same-sex marriage was one of the important concerns of the SGM population. Over the past decade, the United States has made unprecedented progress toward SGM equality and human right. The supreme court passed the same-sex marriage and many states initiated anti-discrimination bills to protect SGM. For example, 23 states in the United States published laws in employment, housing and public accommodation against discrimination based on sexual orientation, gender identity and gender expression to protect SGM people in 2015.²⁷

Nevertheless, not all of the states in the United States have passed their bills. To study the impact of these state-level anti-discrimination laws, we compared the affect processes of SGM people across states with a focus on employment, housing, and public accommodation. From the comparison results, we found that SGM people in New Mexico, which was fully protected by these anti-discrimination laws (i.e. anti-discrimination in employment, housing, and public accommodation), expressed the most positive feeling in their tweets. In contrast, SGM people in Alaska had the highest LIWC scores of negative emotion and anger, which only prohibit discrimination against public employees based on sexual orientation. Arkansas does not have anti-discrimination laws in employment, housing, and public accommodation. Consequently, the SGM Tweeter users in Arkansas had the highest LIWC score of anxiety. Interestingly, even fully protected by these three anti-discrimination laws, the SGMs from Rhode Island had the highest LIWC score of sadness

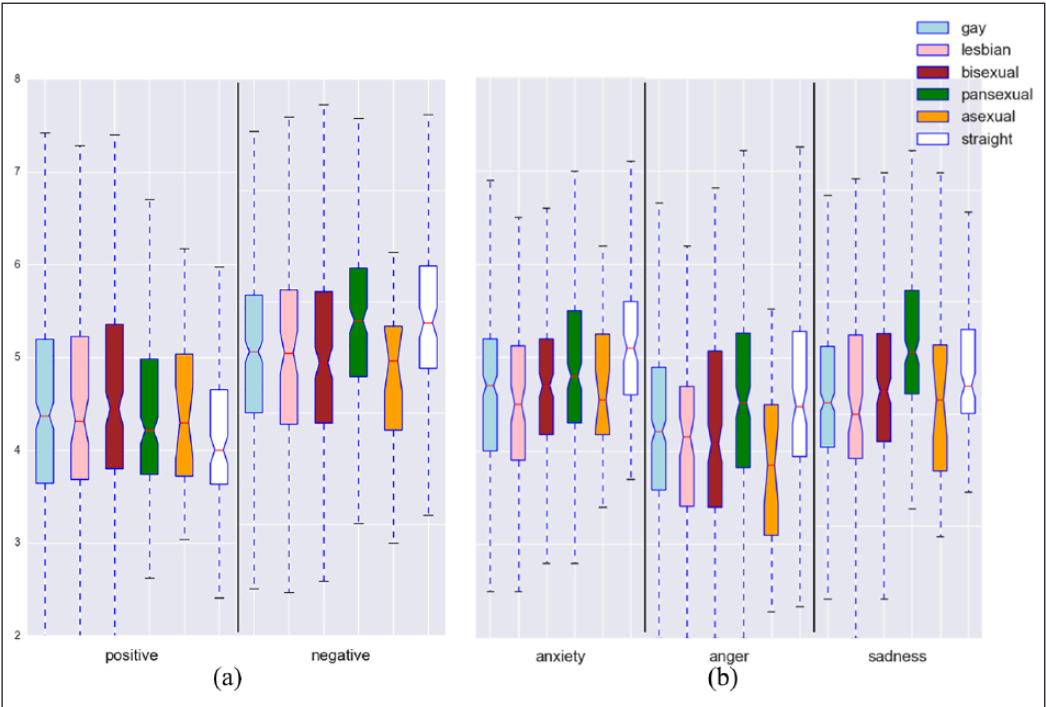


Figure 10. The affect processes ((a): positive vs negative and (b): anger vs anxiety vs sadness) expressed in SGM individuals’ tweets across different SGM groups with different gender identities.

among all the states. A potential explanation is a recent approved bill that bans the chance of reduced sentences or less server punishment for SGM defendant’s violent actions which was called “gay panic” or “trans panic” by the public.²⁸

To investigate which anti-discrimination laws and policies can positively impact the mental health of SGM people, we compared the affect processes of SGM people versus non-SGM people by each state with different anti-discrimination laws and policies (i.e. education, housing, employment, healthcare, hate crimes, public accommodations, convers therapy, and gender on ID).²⁴ In general, compared with the SGM people living in the states with anti-discrimination laws, the SGM people living in the states without anti-discrimination laws expressed more negative feelings. Therefore, we can observe that anti-discrimination laws in general have positive effect on the mental health of the SGM population. However, there is one exception: SGM people living in states with or without explicit “gender on ID policy” showed no statistical differences on their LIWC scores of the affect processes. This observation might indicate that SGM people pay less attention on whether their gender was displayed properly.

Within different sexual orientation subgroups (Figure 2(a)), we found that tweets from pansexual people contained more words related to negative emotion, anger, and sadness than those in other sexual orientation groups (i.e. gay, lesbian, bisexual, and asexual). On the other hand, we noticed that the control group (i.e. straight) used more words related to negative emotion, anxiety than pansexual group. Within different gender identity subgroup (Figure 2(b)), we found that SGMs in the agender subgroup tend to post tweets with words related to negative emotions and anger more often than those in other subgroups (i.e. transman, transwoman, bigender, and genderfluid).

In recent year, social media have emerged as promising platform to examine behavior and attitudes of the minority groups such as SGMs because of their inherent advantages of offering data on variety topics, massive scale and in time.²⁹ In particular, Twitter has a tremendous amount of self-sharing information allow researchers to profile Twitter users by analyzing their posts. Despite these merits, twitter data have several limitations. First, gathering accurate demographic information (e.g. age, gender, race) of users in Twitter is almost impractical. Therefore, further studies of affect processes of SGMs in difference demographic groups are currently infeasible. Second, the results presented in this research may not reflect the truth situation of the whole SGM population since we only studied the SGMs who are Twitter users. Last but not least, the vocabulary using in the social media changing dramatically from year to year. To capture the tweets related to SGMs, we must update the SGMs-related keywords list regularly.

Nevertheless, our study demonstrated the feasibility of using Twitter data as a public health surveillance tool to identify mental health signals in the vulnerable SGM population geographically and at different time period.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding


The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: The study was supported in part by NSF Award #1734134 and NIH UL1TR001427. The content is solely the responsibility of the authors and does not necessarily represent the official views of the sponsors.

Supplemental material

Supplemental material for this article is available online.

ORCID iDs

Yunpeng Zhao  <https://orcid.org/0000-0002-5771-3373>

Jiang Bian  <https://orcid.org/0000-0002-2238-5429>

References

1. Reisner SL, Poteat T, Keatley J, et al. Global health burden and needs of transgender populations: a review. *Lancet* 2016; 388(10042): 412–436.
2. Korda H and Itani Z. Harnessing social media for health promotion and behavior change. *Health Promot Pract* 2013; 14(1): 15–23.
3. Cheston CC, Flickinger TE and Chisolm MS. Social media use in medical education: a systematic review. *Acad Med* 2013; 88: 893–901.
4. Burton-Chase AM, Parker WM, Hennig K, et al. The use of social media to recruit participants with rare conditions: Lynch syndrome as an example. *JMIR Res Protoc* 2017; 6(1): e12.
5. Buckingham L, Becher J, Voytek CD, et al. Going social: success in online recruitment of men who have sex with men for prevention HIV vaccine research. *Vaccine* 2017; 35(27): 3498–3505.
6. Arayasirikul S, Chen Y-H, Jin H, et al. A Web 2.0 and epidemiology mash-up: using respondent-driven sampling in combination with social network site recruitment to reach young transwomen. *AIDS Behav* 2016; 20(6): 1265–1274.
7. Huo H-F and Zhang X-M. Modeling the influence of Twitter in reducing and increasing the spread of influenza epidemics. *SpringerPlus* 2016; 5: 88.

8. Bosley JC, Zhao NW, Hill S, et al. Decoding Twitter: surveillance and trends for cardiac arrest and resuscitation communication. *Resuscitation* 2013; 84(2): 206–212.
9. Jawad M, Abass J, Hariri A, et al. Social media use for public health campaigning in a low resource setting: the case of waterpipe tobacco smoking. *Biomed Res Int* 2015; 2015: 562586.
10. Hanson CL, Cannon B, Burton S, et al. An exploration of social circles and prescription drug abuse through Twitter. *J Med Internet Res* 2013; 15(9): e189.
11. Du J, Tang L, Xiang Y, et al. Public perception analysis of tweets during the 2015 measles outbreak: comparative study using convolutional neural network models. *J Med Internet Res* 2018; 20(7): e236.
12. Hicks A, Hogan WR, Rutherford M, et al. Mining Twitter as a first step toward assessing the adequacy of gender identification terms on intake forms. *AMIA Annu Symp Proc* 2015; 2015: 611–620.
13. Zhao Y, Guo Y, He X, et al. Assessing mental health signals among sexual and gender minorities using Twitter data. In: *2018 IEEE international conference on healthcare informatics workshop (ICHI-W)*, New York, 4–7 June 2018, pp. 51–52. New York: IEEE.
14. Bian J, Yoshigoe K, Hicks A, et al. Mining Twitter to assess the public perception of the “Internet of things.” *PLoS ONE* 2016; 11: e0158450.
15. Bian J, Zhao Y, Salloum RG, et al. Using social media data to understand the impact of promotional information on laypeople’s discussions: a case study of Lynch syndrome. *J Med Internet Res* 2017; 19(12): e414.
16. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011; 12: 2825–2830.
17. Chollet F. Keras. 2015, <https://keras.io> (2018, accessed 13 September 2018).
18. Abadi M, Agarwal A, Barham P, et al. TensorFlow: large-scale machine learning on heterogeneous systems, 2015, <http://download.tensorflow.org/paper/whitepaper2015.pdf>
19. Pennington J, Socher R and Manning CD. GloVe: global vectors for word representation, 2014, pp. 1532–1543, <https://www.aclweb.org/anthology/D14-1162>
20. Wongkoblap A, Vadillo MA and Curcin V. Researching mental health disorders in the era of social media: systematic review. *J Med Internet Res* 2017; 19(6): e228.
21. Xue D and Li F. Research of text categorization model based on random forests. In: *2015 IEEE international conference on computational intelligence & communication technology*, Ghaziabad, India, 13–14 February 2015, pp. 173–176. New York: IEEE.
22. Joachims T. Text categorization with support vector machines: learning with many relevant features. In: Nédellec C and Rouveirol C (eds) *Machine learning: ECML-98*. Berlin; Heidelberg: Springer, 1998, pp. 137–142.
23. Kim Y. Convolutional neural networks for sentence classification. Association for Computational Linguistics, pp. 1746–1751, <https://www.aclweb.org/anthology/D14-1181>
24. Human Rights Campaign (HRC), <https://www.hrc.org/state-maps/> (2018, accessed 13 September 2018).
25. Singh S and Durso LE. Widespread discrimination continues to shape LGBT people’s lives in both subtle and significant ways, 2 May 2017, <https://www.americanprogress.org/issues/lgbt/news/2017/05/02/429529/widespread-discrimination-continues-shape-lgbt-peoples-lives-subtle-significant-ways/>
26. A survey of LGBT Americans. Survey report, <http://www.pewsocialtrends.org/2013/06/13/a-survey-of-lgbt-americans/> (2013, accessed 18 August 2018).
27. Brown TB. Did you know it’s legal in most states to discriminate against LGBT people? NPR.org, <https://www.npr.org/sections/itsallpolitics/2015/04/28/402774189/activists-urge-states-to-protect-the-civil-rights-of-lgbt-people> (2015, accessed 13 September 2018).
28. Rhode Island lawmakers considering bill to ban “gay or trans panic” defense. *Metro Weekly*, <https://www.metroweekly.com/2018/05/rhode-island-lawmakers-bill-ban-gay-trans-panic-defense/> (2018, accessed 13 September 2018).
29. McCormick TH, Lee H, Cesare N, et al. Using Twitter for demographic and social science research: tools for data collection and processing. *Sociol Methods Res* 2017; 46(3): 390–421.