Minorities in Ted Talks channel

Analysis of minority groups in YouTube Ted Talks channel with Text Mining

Benedetta Felici

Course in Computational Social Science

Professors: Christoph Kern & Paul Bauer

GitHub link: https://github.com/bene1f/Computational-Social-Science-24-.git

Abstract

This study uses text mining techniques on comments of TED Talks YouTube videos to investigate sentiment and topics of videos related to minorities. The analysis centres on understanding whether speakers or topics concerning minorities influences comment sentiment and, secondly, on identifying the core themes shaping these responses. Specific minority groups examined include individuals with disabilities, black individuals, women, and members of the LGBTQ+ community. The study employs a mixed-methods approach: it exploits the YouTube API to get the data, sentiment analysis and topic modelling. Key findings indicate a predominance of negative comments on videos related to LGBTQ+ and black people, while videos featuring other minority groups tend to receive more positive responses. Subsequent topic modelling reveals the pervasive presence of existing stereotypes within comments directed toward the LGBTQ+ community and black individuals. This research highlights the potential for online platforms to mirror and show societal sentiment and prejudices towards minorities, emphasising the need to fight hate speech and stereotypes around minority groups.

Contents

1. Introduction and motivation

1.1. Minority representation

In today's diverse and interconnected society, the issue of minority representation has garnered increasing attention. Across various social, political, and cultural spheres, there's a growing concern about whether minority groups receive adequate representation. From media depictions to political engagement, the visibility and inclusion of marginalised communities hold significant implications for societal perceptions, identity formation, and democratic principles. Thus, understanding the scope and quality of minority representation is crucial for addressing systemic inequalities and fostering a more inclusive and equitable society.

Moreover, in the digital age, the importance of minority representation in media is even more pronounced due to the widespread use of media and its influence on public opinion. Numerous studies have examined minority representation across different media platforms, highlighting recurring patterns of insufficient portrayal and stereotypical coverage for women and specific ethnic communities (Schwemmer, C., & Jungkunz, S., 2019). This persistent under-representation not only distorts perceptions among non-minority individuals but also shapes the self-perception of these marginalised groups, perpetuating dominant cultural ideologies and reinforcing societal biases.

One notable study by Merskin (2017) emphasises how mainstream media acts as a powerful tool of hegemony, enforcing dominant narratives regarding race, gender, and socio-economic values. Skewed, inaccurate, or stereotypical representations of societal groups in media shape societal perceptions and influence the way individuals from both minority and non-minority backgrounds perceive certain phenomena.

The study of Schwemmer, C., & Jungkunz, S. (2019) focuses on two different forms of minority representation, namely descriptive and substantive one. Descriptive representation, in digital media, focuses on the presence of minority groups within media platforms, whereas substantive representation evaluates how responsive these platforms are to the needs and interests of marginalised communities.

Scholars suggest that media exposure contributes to the development of stereotypes, negative attitudes, and diminished self-esteem, underscoring the importance of adequate descriptive and substantive representation for women and ethnic groups (Van Klingeren et al., 2015; Martins and Harrison, 2012).

Overall, understanding and addressing issues of minority representation in media are essential steps toward creating a more inclusive and equitable society, where diverse voices are heard and valued. It remains imperative for marginalised groups to have a voice in these spaces to counteract the perpetuation of inequalities.

1.2 Response to minority representation

While much of the existing literature has focused on examining the degree of representation of minorities, both descriptively and substantively, our paper diverges by delving into the response evoked by such representation. We chose to pivot towards examining responsiveness as it offers valuable insights into how these minority groups are perceived and the sentiments they generate.

To understand the power of representation over public opinion we should appeal to the so-called “thermostat model”, according to which public sentiment over minorities is affected by the quality of their representation. In this representation, both their physical presence and their presented topics play an important role. This relationship is exacerbated by media attention (Schwemmer, C., & Jungkunz, S., 2019).

It's crucial to emphasise that this responsiveness not only reflects people's perceptions but also shapes them. Comments from fellow users wield considerable influence over individuals' initial impressions of a video. These social cues can rapidly transform a neutrally presented piece into one perceived as heavily biassed, solely based on the commenting behaviour of viewers (Anderson et al., 2014). For example, on platforms such as YouTube, the tone of comments can exacerbate polarisation or reinforce stereotypes and prejudices against specific social groups or subjects, potentially leading to feedback loops where users align with prevailing opinions (Rothschild & Malhotra, 2014).

1.3. Response to minority representation in Ted Talks Youtube channel

Examining minority representation within the TED Talks YouTube channel presents a fascinating avenue of inquiry. TED Talks serve as a unique platform for disseminating science-related content to a vast online audience. Originating from TED: Technology, Entertainment, and Design conferences established in 1984, these talks are freely distributed across various platforms under the banner of "Ideas Worth Spreading." Since some scholars view TED Talks as a populist source of information rather than exclusively for scientific discourse, their broad accessibility and focus on pressing societal issues make them noteworthy subjects for analysis (Sugimoto and Thelwall, 2013).

The substantial audience also highlights the potential implications of under-representation on the TED stage. Schwemmer and Jungkunz (2019) explored minority representation within TED Talks, focusing on descriptive and substantive aspects while also examining audience responses. Their findings revealed persistent under-representation of non-white speakers, alongside differential sentiment towards talks addressing inequality issues. Particularly noteworthy was the fact that YouTube comments often centred on presenter characteristics; specifically, female speakers elicited stronger emotional responses, both positive and negative.

Inspired by these insights, my research seeks to elucidate how the presence of minorities, whether as speakers or within discussed topics, influences audience sentiment expressed in comments. Through topic modelling of comments, and sentiment analysis, we aim to uncover the underlying factors driving positive and negative reactions to TED Talks.

Our primary research question focuses on assessing the impact of minority presence on comment sentiment, also trying to see if there is sentiment variation between videos with neutral topics but speakers related to minorities and videos with topics explicitly related to minorities. Then, with topic modelling we will analyse recurring themes within, mainly negative, comments to shed light on the reasons for these sentiments.

1.4. Minorities

The term "minority" has evolved beyond simply denoting a small population size. Over time, it came to represent groups, within a hierarchical structure, who hold less power compared to the dominant group. Scholars often use the term "minority group" to describe a demographic group facing relative disadvantages, often linked to their observable characteristics such as ethnicity, race, religion, sexual orientation, or disability.

This study focuses on four specific minority groups: individuals with disabilities, black people, LGBTQ+ individuals and women.
To ensure reliable and accurate sentiment analysis, "control" videos featuring white, male speakers and neutral topics will be included. These control videos act as a baseline for comparison when analysing sentiment towards the other groups.
To explore potential differences in public perception, we will initially categorise the videos into two groups. One group will feature neutral topics but speakers who clearly identify as belonging to a minority group, while the other will encompass videos where the topic explicitly relates to a minority, as indicated by the title as well. By analysing comment sentiment, we aim to identify variations in public perception between these two categories and among the different minority groups.

2. Data and Methods

2.1 Data

The data are taken from YouTube platform through YouTube API, the main core of the dataset are the comments of each video. Given the overwhelming number of comments on the TED Talks channel, conducting a comprehensive analysis of all comments would have been unfeasible. To address this challenge, we implemented a strategic selection approach. Initially, we concentrated on videos specifically addressing minority groups, including disability, race, sexual orientation, and gender. Within this subset, we prioritised videos with the highest comment volumes. This method yielded a representative dataset comprising 116,994 comments, which formed the basis for our analysis. In the following section we will see how exactly these data were obtained and how they have been analysed.

2.2. Methods

2.3. Youtube api

To access data from YouTube, the official YouTube API was utilised. This API requires a personal key and allows authorised users to retrieve information about videos and channels, including information on comments, likes, and views.
For this task, we employed the “tuber” package in R, which facilitates interaction with the YouTube API. This tool was fundamental to gather the data necessary for subsequent analyses.

2.4. Sentiment Analysis

Sentiment analysis serves as a valuable technique for evaluating opinions, sentiments, and emotions conveyed in text. When applied to individual words, it entails assigning sentiments based on specific dictionaries, such as Bing, which provide sentiment and scores indicating positivity or negativity. Conversely, when applied to entire comments, sentiment analysis typically assigns a polarity to the comment based on its overall score, calculated by summing the scores of individual words. This score is then adjusted to account for valence shifters like "not" or "really," as well as for word amplifiers.

However, this method has some limitations. First, the frequent use of emojis poses a challenge since they cannot be directly interpreted by the majority of algorithms. We should transform them to avoid misinterpretation of sentences and to avoid losing important information. Another limitation lies in the lack of context awareness, analysing individual words without considering their context can lead to misjudgement.

To mitigate these limitations, this analysis will undertake several steps.
First of all, we will transform emojis into their associated emotions, ensuring their contribution to sentence meaning is accurately captured.
Secondly, we will use two analysis methods, one based on the analysis of individual words and another based on the analysis of the entire comment.
In the first case, each word will be categorised as positive or negative using sentiment dictionaries, specifically Bing and the "tm" package in R.
In the second case we will consider the entire comment for a better idea of the context, taking into account valence shifters for a more accurate interpretation. Each comment will be assigned a polarity (positive or negative) based on its overall sentiment score, facilitated by the "qdap" package in R.

2.5. Topic modelling

Topic modelling, a method for automated content analysis, allows us to uncover hidden thematic structures within large collections of text, comments in our case. Understanding prevalent themes within comments is crucial for deciphering the underlying reasons behind positive or negative sentiment and for discerning which specific aspects evoke these reactions.

To prepare the textual data for topic modelling, we utilised the R programming language and its associated packages, primarily the package “quanteda”. These tools facilitated the conversion of comments into a structured corpus, suitable for further analysis. Notably, punctuation, symbols, numbers, and stop words were removed during the preparation process to select only relevant words. In this case, the outputs are tables divided in rows, where each row represents a specific discovered topic and contains the words associated with it.

3. Results

My main analysis focuses on two distinct representations of minorities: through the speaker and through the topic of the video.

In the first part, we analyse videos with neutral topics featuring speakers from two specific minority groups: black women and individuals with disabilities. This allows us to understand potential differences in sentiment based on the speaker's identity.

The second part examines videos where the topic is explicitly, from the title, related to specific minorities. These include the LGBTQ+ community (with an additional focus on transgender individuals), women, and black people. This approach allows us to explore potential shifts in sentiment based on the video's subject matter.

The overarching goal is to examine whether the presence of minorities, either through the speaker or the topic, influences the sentiment expressed in the comments compared to a control group. This control group consists of videos with white, male speakers addressing neutral topics.

To achieve this objective, I employ a multi-step process:

- Sentiment analysis on single words: Assessing the sentiment within each minority group expressed by individual words.

- Sentiment analysis on full comments: Assessing the sentiment within each minority group expressed by polarity of entire comments.

- Topic modelling: Assessing the underlying themes within each minority group's comments.

3.1 Sentiment analysis on single words

My initial analysis focused on videos with neutral topics featuring speakers from minority groups. As illustrated in Fig. 1, the distribution of positive and negative words across these groups suggests a general trend of positivity. Infact, for each group, namely control, black women and disability group, we see, both from the graph and analytically, that there are more positive than negative words. Specifically, we find that 70% of words are positive for the control group, 66% for the disability group and 85% for black women group.

The graph in Fig.2 focused on videos featuring, explicitly from the title, topics related to minority groups.
Interestingly, as you can see in Fig.2, in this case, sentiment analysis reveals contrasting trends. For LGBTQ+, transgender, and black people, the distribution of positive and negative words shows a higher proportion of negative words. Specifically, we found that 61% of words are negative for transgender group, 60.5 % of words for black people group and 57% for LGBTQ+ group. Conversely, the control group and videos related to women exhibit more positive than negative words, 70% of words are positive for the control group and 53% for the women group. To corroborate our findings we will look at the polarity of the entire comments to see if the results will confirm or disconfirm these initial findings.

3.2. Sentiment Analysis on entire comments

Shifting the focus from analysing individual words to entire comments allowed us to understand the surrounding context and, in turn, a more accurate understanding of the sentiment.

For the first category, namely videos with the speaker related to minority groups, Fig. 3 confirms the trend observed previously. We can notice a positive overall polarity of comments across all groups, indicated by the rightward shifts on the graphs. So, we can underline that the presence of a speaker belonging to a minority group does not seem to affect the sentiment in the comments, given that all the three groups of videos present more positive than negative comments.

For the second category, namely videos with the topic related to minority, Fig. 4, further corroborates the findings from the previous analysis. Consistent with the observations based on individual words, groups related to LGBTQ+, transgender individuals, and Black people exhibit a slight leftward shift in polarity, indicating a greater prevalence of negative comments. Conversely, the control group and the women's group maintain a rightward shift, suggesting a dominant presence of positive comments.

This consistency between the single-word and entire-comment analyses strengthens the conclusions drawn about the influence of topics on comment sentiment. While the speaker's identity appears not to significantly impact sentiment, the video's topic seems to play a more influential role, specifically, videos with topics related to LGBTQ+ group and to black people present more negative than positive sentiments.

3.3. Topic modelling

To delve deeper into comments content and potentially uncover reasons for negative sentiment, as well as the subtopics triggering such reactions, we utilised topic modelling. We applied this technique to the comments associated with the groups that receive predominantly negative comments, namely LGBTQ+ group, transgender group and black people group.

3.3.1. LGBTQ+ Group:

Following topic modelling, several themes emerge within this minority group, each characterised by distinct keywords (Fig. 5). It is crucial to interpret these findings through a human lens, not solely relying on the algorithm's output.

For example, words like "Jesus," "Christ," "God," and "Satan" suggest that religion plays a significant role in discussions surrounding the LGBTQ+ community. Additionally, terms such as "vomiting," "nauseated," "poo," "horrible," and "crap" appear to relate to the topic of disgust. Furthermore, words like "immoral," "predator," "pedo," and "indoctrination" raise potential concerns about morality in viewers' minds, possibly linking the LGBTQ+ community with fears regarding public safety. Additionally, words like "kink," "hate," "sick," "dislike," "doctor," "kill," "worry," "bullshit," "wtf," and "bs" capture a general sense of hate and negativity. It is important to acknowledge that a small number of positive words were also identified, including "joy," "tolerance," and "love."

The analysis of specific words and themes, within the comments, allows for a deeper understanding of deeply ingrained stereotypes and prejudices associated with the LGBTQ+ community. Notably, certain words identified through topic modelling exhibit a strong correlation with existing societal biases against this group.

Firstly, words related to religion suggest a specific stereotype: the notion of homosexuality being inherently sinful or ungodly. This perspective often stems from interpretations of specific scriptural passages, such as those found in Leviticus, Corinthians or in the story of Sodom and Gomorrah (Gnuse, Robert K., 2015).

Secondly, the appearance of words pertaining to "indoctrination" leads to the misconception of "homosexual recruitment." This harmful allegation, suggests that LGBTQ+ individuals intentionally engage in concerted efforts to influence children into adopting a homosexual identity (Posner Sarah, 9 February 2007).

Furthermore, the presence of terms like "pedo" or "predator" highlights another dangerous stereotype, linking gay men with sexual predation and pedophilia. This harmful association has been consistently refuted by extensive research (Dynes, Wayne R., 2016).

Finally, the emergence of words like "sick" or "doctor" reflects another stereotype: the fallacy of homosexuality as a disease. This erroneous perspective has a long and harmful history, with instances like Nazi propaganda portraying homosexuality as a contagious threat to the nation (Giles, Geoffrey J., 2010).

3.3.2. Transgender group:

Topic modelling allows us to identify recurring themes within comments related to the transgender community. However, careful consideration is needed when interpreting the results.

The presented keywords in Fig. 6, highlight various topics present in the comments.
First of all, words like “hitlery”, “libtard”, “left”, “revolution”, "laws," "justice," and "vote" potentially point towards politics and legislation. This might be influenced by contents

related to gender-neutral bathrooms, which are often subject to regulation and political debate. Then, there are several words related to hate, like: “awkward”,”suicide”, “fuckin”, “stupid”, “sick”, “idiots”, “bs” and “dislike”. Furthermore, words like "emergencies," "abuse," and "problems" hint at potential concerns regarding public safety, as in the example of LGBTQ+ community.

Also in this case topic modelling allows for a better awareness of rooted and spread stereotypes associated with the transgender community. Specific keywords are strongly related to existing societal biases against this group.

The comments reveal a strong focus on legislation and bathroom access, which directly relates to the harmful "bathroom predator" stereotype surrounding transgender individuals. This stereotype frequently manifests in proposed "bathroom bills", which seek to deny transgender individuals access to public restrooms aligned with their gender identity (Kasperkevic, Jana, March 17, 2015).
Proponents of such bills often justify this legislation based on claims of maintaining privacy, preventing assault and rape. Additionally, they argue that these restrictions are necessary for psychological comfort (Bianco, Marcie, April 2, 2016).
Critics of this legislation strongly refute these arguments, contending that bathroom bills do not enhance the safety of cisgender individuals. Furthermore, they assert that such policies actively endanger both transgender individuals and gender-nonconforming cisgender individuals (King-Miller, Lindsay, April 12, 2016).

3.3.3. Black People

After topic modelling applied to the videos related to black people group, certain words stand out, such as "gene," "evolution," "biological," and "race," which have historically been used by racists to justify racism by using incorrect biological concepts. Additionally, terms like "yellow," "Aryans," "Slavic," and "Ainu" are associated with the notion of race. There's also a vocabulary tied to violence, including words like "bloodshed," "guilt," "kidnapping," and "annihilation." Also the lexicon of slavery is widely present, encompassing words like "enslave," "slavery," "slaves," "tribery," "colonials,". Furthermore, there are words linked to American politics, such as "Trump" and "Democrats." Finally, there are terms concerning human rights, like"Black Lives Matter," "minorities," and "rights."

Certain words and themes within comments highlight the harmful presence of stereotypes, revealing their deep-seated nature within society.

First of all, the character of Jim Crow, which embodies all the stereotypes about black people, was present in several comments (Rehin, George F., December 1975).

Moreover, Words like "guilt" and "kidnapping" reveal the stereotypes of the 1980s and subsequent decades, depicting Black men as violent and dangerous criminals, contributing to a persistent negative perception (Drummond, 1990).
The pervasiveness of these societal biases is further exemplified by the existence of the racial slur "Black Buck," which historically caricatured Black men as resistant to authority and inherently violent (Oliver, 2003).

Discussions surrounding slavery not only address the historical suffering endured by African Americans but also draw attention to the "mammy stereotype." This historical U.S. stereotype, as described by Walker-Barnes (2014), depicted Black women, primarily enslaved, as domestic workers responsible for child care, perpetuating harmful generalisations.

These examples, related to the words and topics found in the comments, demonstrate the enduring presence of damaging stereotypes and prejudices within society, highlighting the need for continued efforts to dismantle these pervasive and harmful social biases.

4. Limitations

The study encounters several limitations that impact its comprehensiveness and the interpretation of its findings. Firstly, relying solely on the YouTube API restricts access to crucial speaker data like gender and ethnicity. These would have opened more research questions and made the analysis much more significant.

Then, dividing the analysis into speakers' and topics' sections poses a challenge when these aspects are heavily intertwined. In cases where the speaker and the topics they discuss are closely related, isolating the individual effects of each becomes difficult. To achieve this separation, the study would require access to wider speaker information and employ regression analysis.

Moreover, the two methods used, namely sentiment analysis and topic modelling, part of a wider method that is text mining, have inherent limitations that will now be discussed furtherly.

Firstly, the lack of context in text analysis poses a challenge. Analysing text snippets like words without considering the broader context in which they were written can lead to misinterpretations. This applies both to sentiment analysis, where the meaning and intent behind words are crucial, and to topic modelling, where identifying the core theme requires understanding the surrounding context. While sentiment analysis utilises polarity analysis to mitigate this issue to some extent, with topic modelling it is more difficult. In fact, short comments often lack the necessary context for accurate interpretation, potentially leading to misconception by the text mining algorithms. Additionally, the limited length of comments might not provide enough information for the algorithms to identify topics with precise accuracy, potentially resulting in overly broad or irrelevant themes being identified. Furthermore, these algorithms often struggle with the informal language commonly found in online comments, such as slang, sarcasm and non-standard language. This can hinder their ability to accurately extract the intended meaning from the text. Finally, topic modelling algorithms focus on identifying recurring word patterns, this could lead to the creation of topics based on frequency of words instead of their meaning. This could result in a superficial and mechanical creation of topics, for this reason a human eye is more than fundamental.

5. Discussion and Conclusion

This research delves into a relatively unexplored area, examining people's responses to videos related to minorities on TED Talks. Our analysis unveils a multifaceted interplay among video topics, speaker identities, and comment sentiment on TED Talks. Notably, comments on videos related to LGBTQ+ and Black communities exhibited a greater degree of negativity. Specifically, we find that 61% of words express negativity for the transgender group, 60.5% for the black community, and 57% for the LGBTQ+ group. In contrast, videos regarding women and featuring minority speakers in neutral contexts received primarily positive comments. Specifically, 53% of words express positivity for the women group, 66% for the disability group, and 85% for the black women group. Notably, for videos with white, male, cisgender speakers talking about neutral topics, the percentage of positive words was 70%.

These findings are supported by the analysis of polarity within comments, where topics related LGBTQ+ and black people are associated with negative polarity. On the other hand, videos related to women and videos with neutral topics featuring a minority speaker are associated with a positive polarity.

For the three minorities with more negative than positive comments (namely LGBTQ+ and black communities), we conducted topic modelling to identify the subtopics associated with these negative sentiments. The results revealed that main themes revolved around existing stereotypes. For instance, for the LGBTQ+ community, words related to religion and diseases emerged, while for black communities, words associated with criminality and slavery.

Since comments not only reflect but also actively shape perceptions, it is imperative to be vigilant regarding their contents and take proactive measures to combat prejudices and stereotypes. This underscores the necessity of taking actions to prevent the perpetuation of social biases and hate speech.

Addressing the root causes of online negativity and fighting the spread of harmful stereotypes requires a multi-pronged approach. Firstly, promoting media literacy empowers individuals to critically analyse online content. Secondly, implementing stricter regulations and effective reporting mechanisms can deter online harassment and hate speech. Moreover, fostering open dialogue and challenging negative biases through educational initiatives and public discourse are crucial steps towards creating a more inclusive and respectful online environment for all.

6. Appendix

6.1 Plots

Fig. 1. Sentiment Analysis on words, groups of speakers

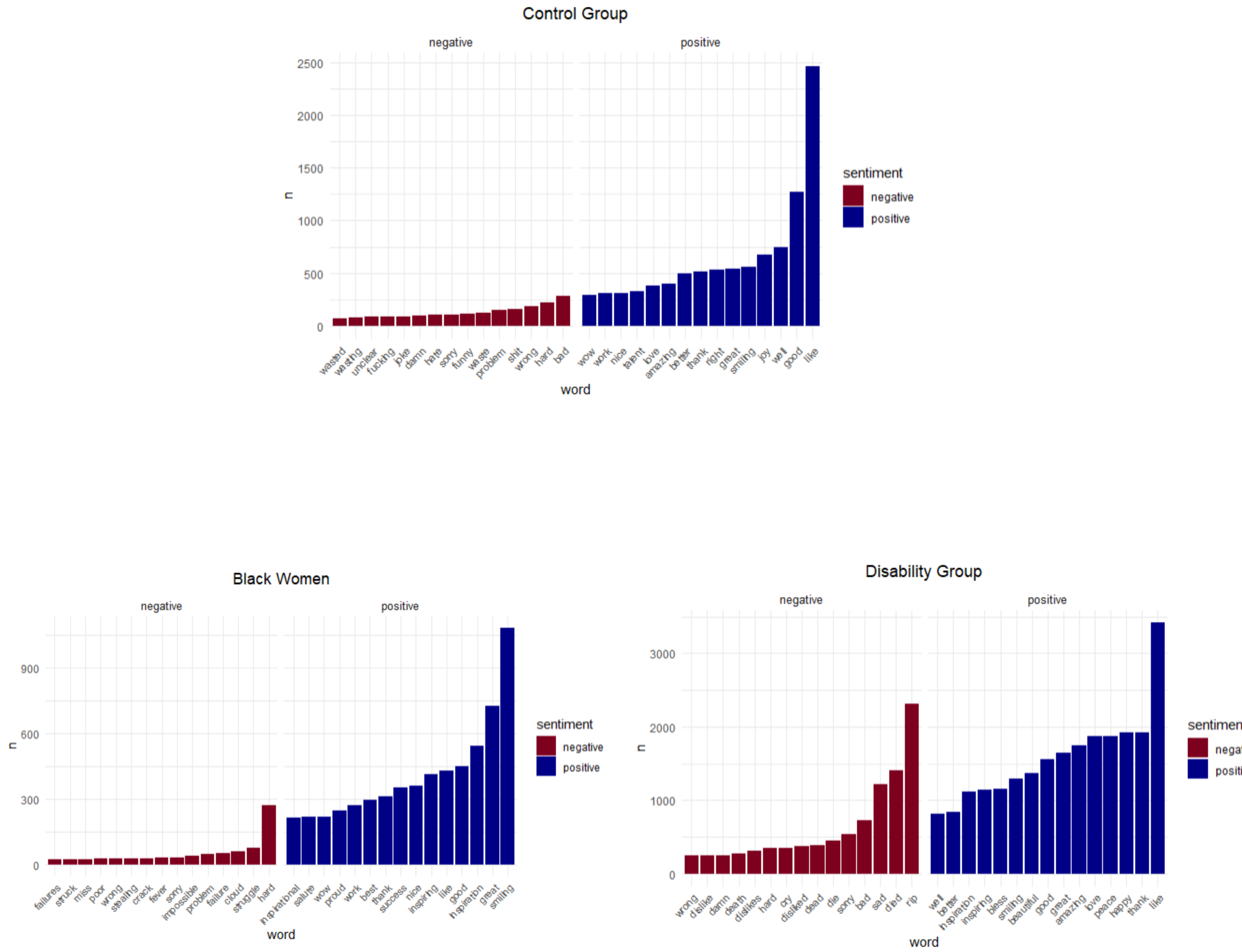

Fig. 2 Sentiment Analysis on words, groups of topics

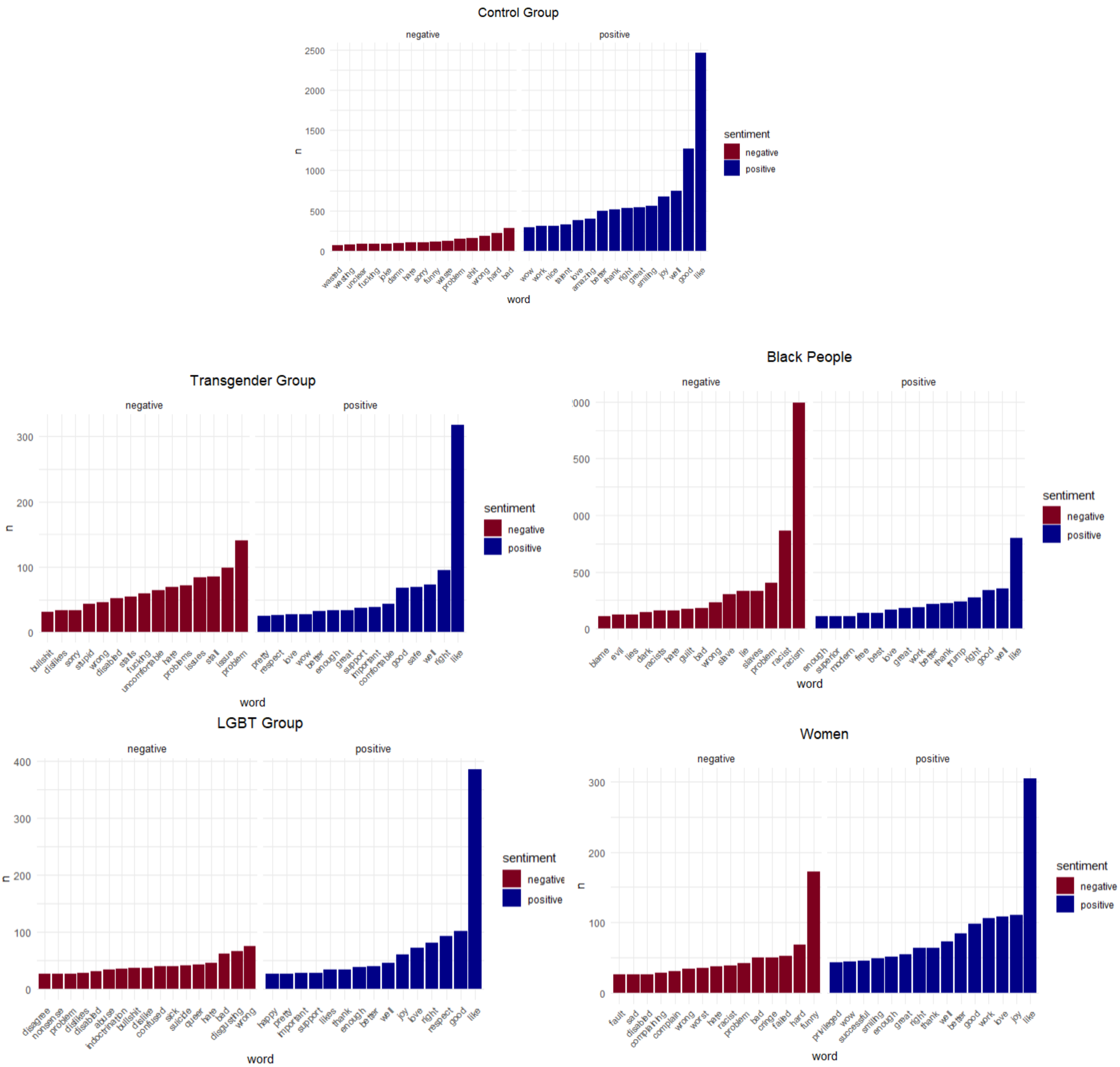

Fig. 3: Sentiment Analysis on comments, groups of speakers

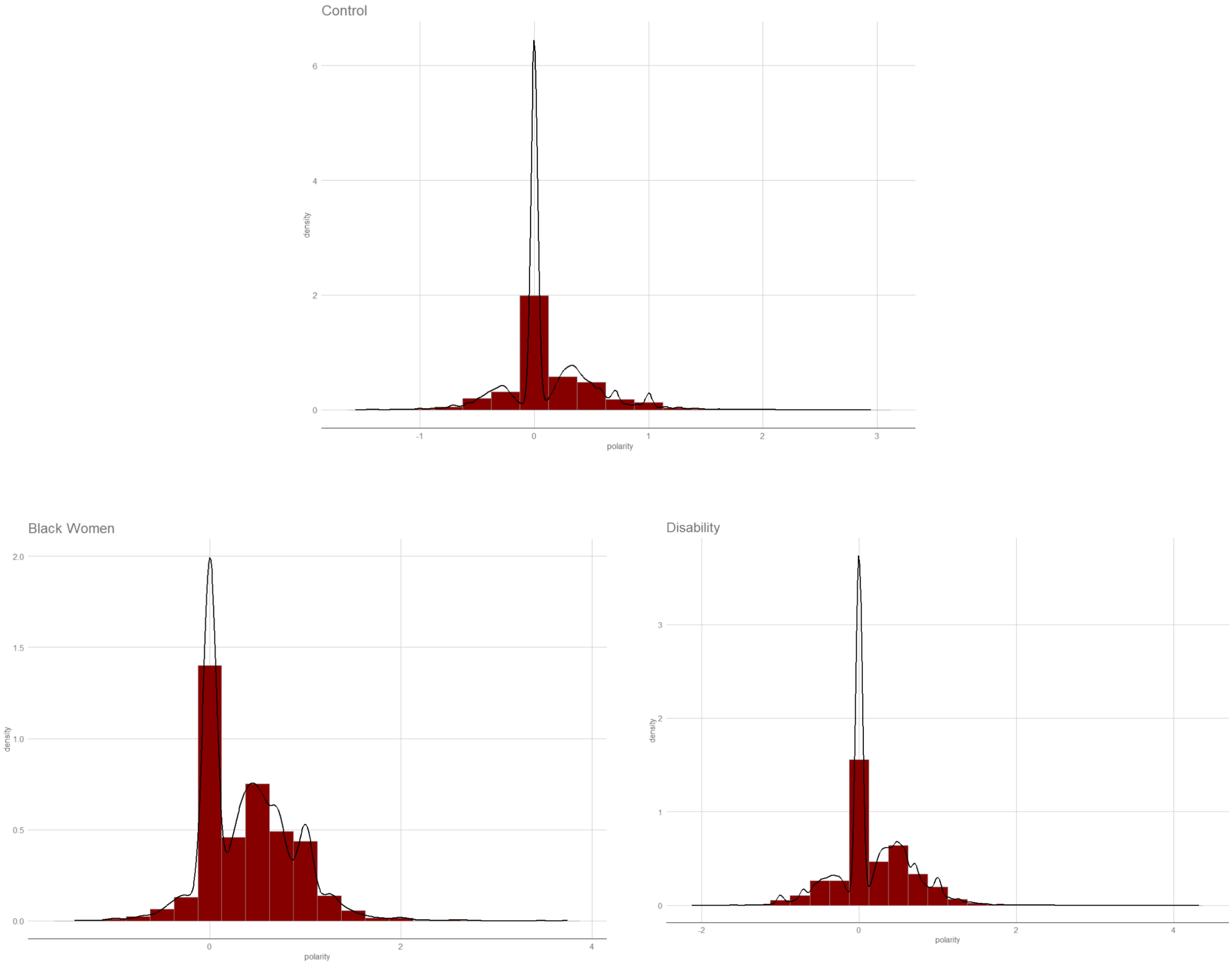

Fig. 4: Sentiment Analysis on comments, groups of topics

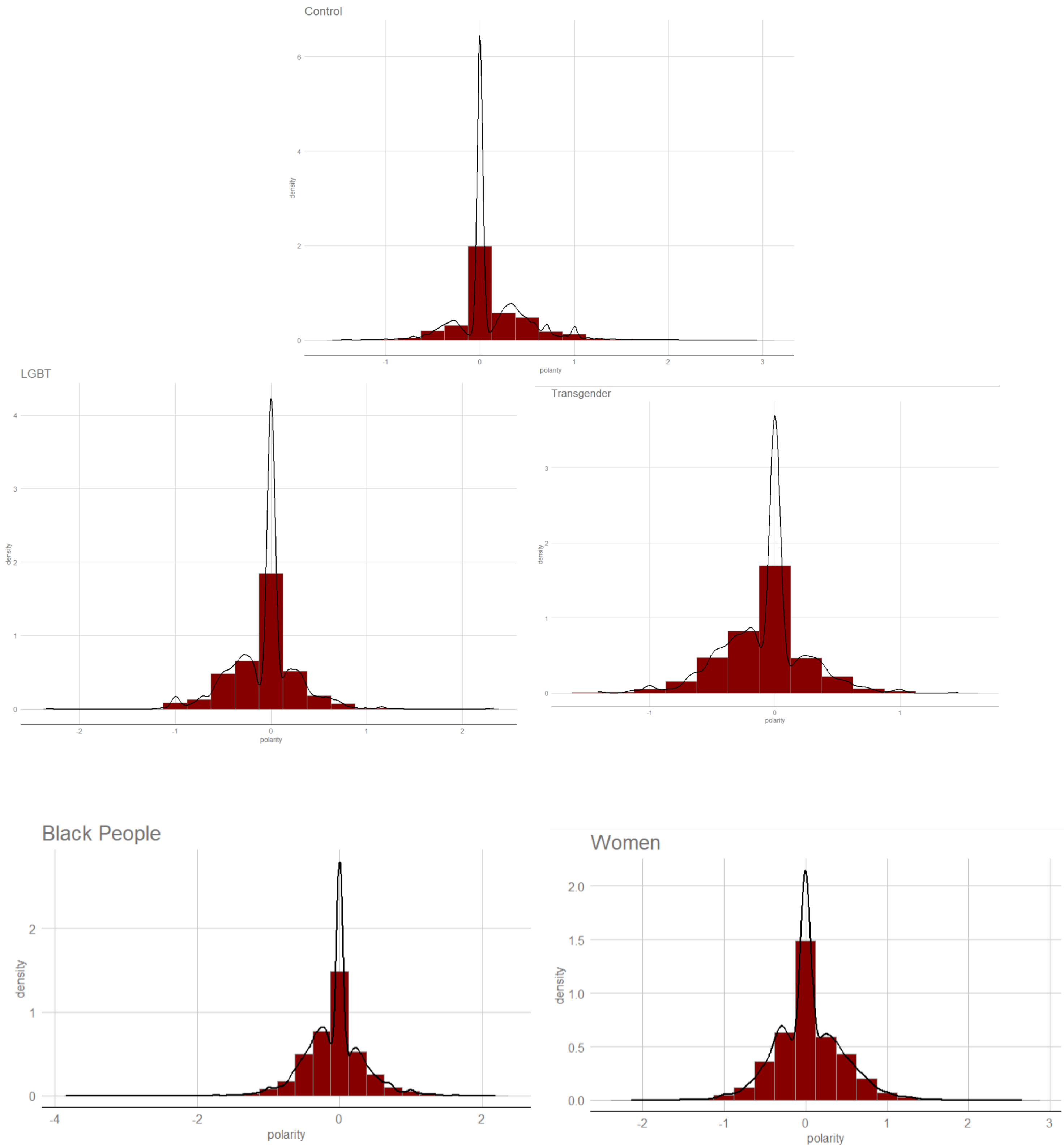

Fig. 5. Topic modelling LGBTQ+ group

Topic 1:
leave, joke, let, need, kids, sexual, thats, person, actually, bullshit, learn, away, thing, figure, alone, predator, sexuality, crap, adults, kink

Topic 2:
wtf, anywhere, bs, jail, jesus, near, piece, satan, christ, humans, interested, looks, mediocre, talk, white, hate, sick, lot, sign, arabic

Topic 3:
face, tears, vomiting, ukulele, fire, hibiscus, nauseated, rolling, blossom, fucked, medium, pile, poo, god, develop, lets, roles, smiling, joy, best

Topic 4:
might, calling, able, speech, creates, minds, decision, attracted, they're, telling, everything, use, yes, answer, indoctrination, taken, explain, around, don't, young

Topic 5:
dislike, ratio, section, cringe, respected, alphabet, spreading, acknowledged, believes, horrible, worth, forget, la, que, whining, video, pedo, oh, comment, dislikes

Topic 6:
comments, thank, get, balls, haters, tolerance, love, lgbt, us, keep, speak, woke, confused, stop, queer, try, start, trying, classes, fragile

Topic 7:
doctor, thumbs, toys, girl, well, want, i'm, road, needs, like, boy, isn't, daughter, uncomfortable, raise, group, wish, indoctrinate, instead, kill

Topic 8:
middle, phobia, she's, bi, gay, youtube, bringing, high, nobody, concepts, school, doesn't, appropriate, thanks, parent, amer's, colours, immoral, rate, rates

Topic 9:
question, relationships, helping, law, understand, type, things, important, beliefs, given, worry, taught, live, pink, learning, generations, newer, princesses, ages, define

Fig 6: Topic modelling Transgender group

Topic 1:
accurate, brush, urinals, men, gender-neutral, bathrooms, will, common, women, often, small, awkward, drops, separated, still, building, hair, apply, areas, laws

Topic 2:
preparedness, political, generate, plan, justice, suicide, channel, gave, remember, speakers, section, brains, carlos, emergencies, fire, fuckin, muslims, particularly, etc, gets

Topic 3:
excellent, toilets, mind, hitlery, libtard, rotted, rooms, confused, wait, stupid, individuals, need, lauri, raudsepp, scared, stopped, wherever, dont, vote, belong

Topic 4:
sick, tears, joy, grinning, idiots, face, eyes, today, labels, ppl, left, add, everything, came, time, smiling, ok, discredit, we're, mean

Topic 5:
abuse, seriously, tedx, victims, piss, revolution, second, versa, vice, tell, half, sexual, males, bs, drunk, ill, legally, trigger, washroom, saying

Topic 6:
valid, looks, realize, camoulflage, ratio, traditionally, dislike, even, children, genders, restroom, another, addition, really, currently, thought, sounds, fit, show, good

Topic 7:
problems, tiny, people, yeah, help, ask, wow, pretty, say, bigger, gonna, chose, suck, black, well, disagree, thinking, whole, solve, normal

Topic 8:
exist, schools, female, male, toilet, medical, single, stall, non, it's, one, work, decades, ellabunny, talked, give, ladies, disabled, dude, sorry

Topic 9:
risk, boy, vagina, men's, boys, grade, u, neutral, bathroom, available, define, spaces, wheelchair, based, ever, deal, lol, love, women's, tried

Fig.10 Topic modelling Black people group

Topic 1:
face, eyes, rolling, joy, smiling, students, floor, laughing, youtube, journalist, collision, town, thanks, bodies, report, democrats, watch, tears, speech, video

Topic 2:
answer, accurate, problem, nature, humans, solve, responsibility, natural, truly, power, half, lie, evolution, function, reverse, w, unfortunately, divide, big, however

Topic 3:
medium-dark, tone, thank, 10th, person's, bloodshed, danish, denmark's, hundred, nice, dark, u, hope, upon, bless, gene, thumbs, foreign, book, raising

Topic 4:
relations, manliness, crt, intellect, aryans, vir, race, plain, slavic, ainu, utter, scientific, critical, characteristics, assume, construct, english, biological, habits, greeks

Topic 5:
color, white, people, person, ppl, it's, guilt, millions, group, responsible, skin, something, lives, issue, matter, don't, thing, worse, stupid, yellow

Topic 6:
africans, sold, enslave, kidnapped, traders, kidnap, advanced, asia, arab, europeans, travel, slavery, african, colonial, tribes, slave, africa, slaves, waste, portuguese

Topic 7:
de, que, pike, o, aryan, albert, lectures, os, wood, data, col, em, majesty, blind, mais, não, annihilation, council, discovery, greatness

Topic 8:
apply, sounds, talk, speak, just, mainstream, signal, treat, willing, system, stand, fine, liberals, anyway, appropriate, booker, caring, save, voice, used

Topic 9:
blacks, neighborhoods, folks, minorities, racist, blm, bad, trump, free, racists, whites, minority, america, care, crow, jim, realize, business, workers, right

References

1. Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2014). The “nasty effect:” Online incivility and risk perceptions of emerging technologies. Journal of computer-mediated communication, 19(3), 373-387.
2. Bianco, M. (April 2, 2016). Statistics Show Exactly How Many Times Trans People Have Attacked You in Bathrooms. Retrieved January 11, 2017.
3. Drummond, W. J. (1990). About Face: From Alliance to Alienation. Blacks and the News Media. *The American Enterprise, 1*(4), 22–29.
4. Dynes, W. R. (2016). Myths & Fabrications. In *Encyclopedia of Homosexuality: Volume II* (p. 870). Routledge.
5. Giles, G. J. (2010). The Persecution of Gay Men and Lesbians During the Third Reich. In *The Routledge History of the Holocaust* (pp. 385–396). Routledge.
6. Gnuse, R. K. (2015). Seven Gay Texts: Biblical Passages Used to Condemn Homosexuality.
7. Harp, D., Bachmann, I., & Loke, J. (2014). Where are the women? The presence of female columnists in US opinion pages. Journalism & Mass Communication Quarterly, 91(2), 289-307.
8. Merskin, D. (2017). Media Representation: Minorities. The international encyclopedia of media effects, 1-10.
9. Kasperkevic, J. (March 17, 2015). Florida anti-transgender bathroom bill moves a step closer to passing. *The Guardian*. Retrieved July 11, 2017.
10. King-Miller, L. (April 12, 2016). Stop Using Women's Safety to Justify Transphobia. Retrieved January 11, 2017.
11. Posner, S. (9 February 2007). The gay recruit. *Chicago Sun-Times*. Retrieved May 2, 2012.
12. Rehin, G. F. (December 1975). Harlequin Jim Crow: Continuity and Convergence in Blackface Clowning. *The Journal of Popular Culture, 9*(3), 682–701.
13. Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder‐Luis, J., Gadarian, S. K., ... & Rand, D. G. (2014). Structural topic models for open‐ended survey responses. American journal of political science, 58(4), 1064-1082.
14. Rothschild, D., & Malhotra, N. (2014). Are public opinion polls self-fulfilling prophecies?. Research & Politics, 1(2), 2053168014547667.
15. Schwemmer, C., & Jungkunz, S. (2019). Whose ideas are worth spreading? The representation of women and ethnic groups in TED talks. Political Research Exchange, 1(1), 1-23.
16. Sugimoto, C. R., & Thelwall, M. (2013). Scholars on soap boxes: Science communication and dissemination in TED videos. *Journal of the American society for information science and technology*, *64*(4), 663-674.
17. White, D. G. (1999). *Ar'n't I a woman? : Female slaves in the plantation South*. New York: W.W. Norton.