

Homophobia and Misogyny in Rap and Hip Hop

By Ankush Bharadwaj

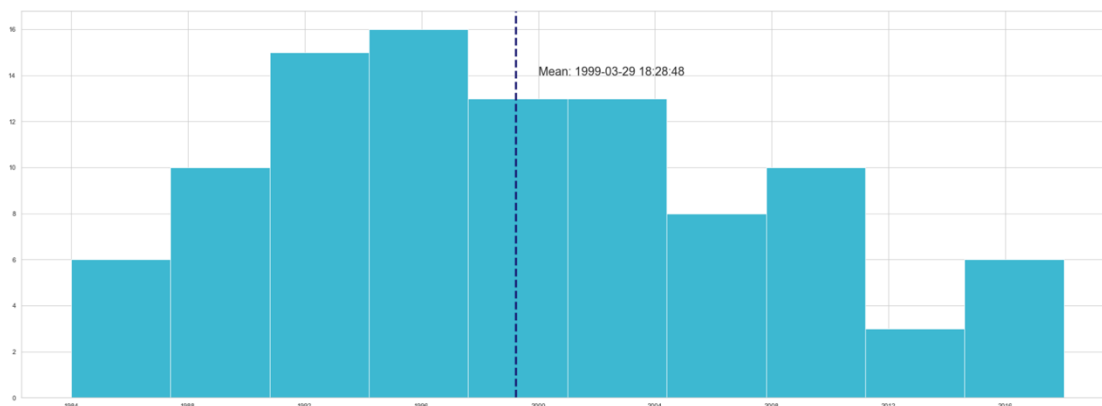
Introduction

Controversy seems to follow rap and hip hop, as the genre has been repeatedly called out for portraying homophobic and misogynistic themes in its music^{1,2}. The use of profanity in the genre is widespread, and unfortunately, the profane content of many rap and hip hop songs includes demeaning references and violent lyrical jabs at women². Furthermore, even the best MCs shy away from being misinterpreted as gay, because “you can’t be fuckin’ people in the ass and say you’re gangsta”, as Method Man, famous rapper and member of the East Coast hip hop collective Wu-Tang Clan said¹. Needless to say, the integration of homophobic and misogynistic themes within rap and hip hop music is worrying, especially given that this genre is among the most popular genres young people listen to, causing concern about the impact of these themes on susceptible minds. However, not all rappers are homophobic and misogynistic. Likewise, not every song or album in the genre entertains homophobic and misogynistic themes. Through this project, changes in the expression of homophobic and misogynistic themes in rap and hip hop albums over time will be studied. To do this, the use of homophobic and misogynistic slurs in 100 rap and hip hop albums ranging from 1984 to 2018 will be analyzed.

Process

The code to complete this project was put together in a Jupyter Notebook using Python 3.

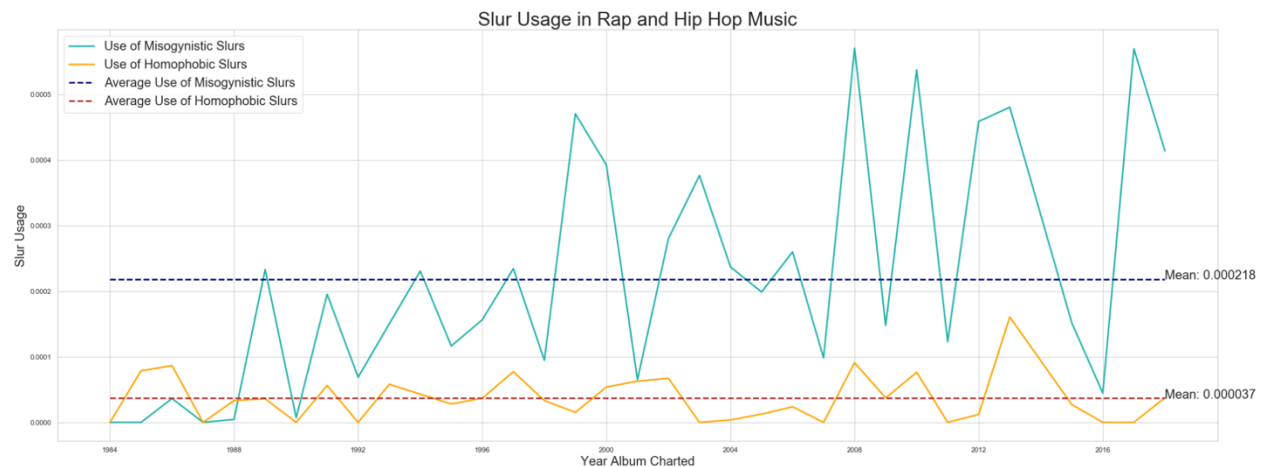
To accomplish this project, a great deal of web scraping is required. First, a list of albums to study will be required. The list at this [link](#) includes the 100 top rap albums and was aggregated from twenty-five other lists, which hopefully means that the songs on this list are an adequate encapsulation of the genre’s standard on the year the album had charted. After web scraping the list of albums into a Pandas DataFrame, I grew worried that perhaps the albums weren’t as well-distributed over time as I would like. Chiefly, I was worried that there would be an overemphasis placed on older albums. To quench my anxiety, I found that, although the distribution does skew slightly right, with most of the albums charting before of the average year across the DataFrame, I shouldn’t be too worried about having unequal representation in terms of the year the album charted in my dataset:



The next step was to scrape the track lists for each album and the lyrics for each track within an album: genius.com was used to acquire the track lists of each album, as well as the lyrics of each song in each album. Unfortunately, at this step I was faced with the seemingly impossible task of assembling a list of some 1,863 URLs to scrape the lyrics of each song. Luckily, by noticing that the URLs used to access genius.com have a predictable pattern, I efficiently compiled the lyrics of most of the songs in the DataFrame, although I unfortunately lost about 7.24% of the original songs in the DataFrame due to errors in my attempt to fully automate the web scraper for the lyrics. Now that the lyrics for each of the songs was in the DataFrame, I could continue to perform text analysis.

The first step of my text analysis methodology was to define the list of homophobic and misogynistic slurs to search for. The list of homophobic slurs was defined to include *sissy*, *pansy*, *sodomite*, *nancy*, *fairy*, *faggot*, *bent*, *fag*, *jockey*, and *tonk*. The list of misogynistic slurs was defined to include *whore*, *slut*, *skank*, *pussy*, *cunt*, *dyke*, *bitch*, *thot*, *prude*, and *lesbo*. Next, to determine how many misogynistic and homophobic slurs and album contained, I first found the TF-IDF scores for each of the slurs in each list within each album. Then, for each overall album, the score given for homophobia and misogyny would simply be the averaged TF-IDF scores for each of the slurs in the respective list. By doing this, I attempted to capture how important misogynistic and homophobic words were to specific albums relative to the use of the same words in all the other albums within the DataFrame. Armed with this information, it was finally time to formulate my results.

Results



The above chart demonstrates the average score for homophobia and misogyny for all rap and hip hop albums in my dataset that charted in a specific year. Although it's difficult to determine whether or not the use of homophobic slurs has changed at all over time, it's very apparent that the usage of misogynistic slurs over time has increased. Notice that it wasn't until the late 1990s that albums began using more misogynistic slurs than the average number of misogynistic slurs released over all albums in the dataset. From this point on especially, it seems as though there's more variation in the use of misogynistic slurs in hip hop and rap music released each year compared to the use of homophobic slurs. After the standard deviation was calculated, this was found to be true, as the standard deviation for the usage of misogynistic slurs is nearly five times higher than the standard deviation for the usage of homophobic slurs.

Furthermore, my analysis found that the most homophobic artists in the dataset are as follows: Snoop Dogg (or, as he was known at the time of the release of *Doggystyle*, his only album in the dataset, Snoop Doggy Dogg), Missy Elliott, and Eminem. My analysis also found that the most misogynistic artists in the dataset are as follows: Lil Wayne, D12, and Post Malone. Similarly, my analysis found that the three most homophobic are as follows: *Doggystyle* by Snoop Dogg, *808's and Heartbreak* by Kanye West, *Raising Hell* by Run-D.M.C. Finally, the most misogynistic albums are as follows: *Tha Carter III* by Lil Wayne, *2001* by Dr. Dre, and *The Marshall Mathers LP* by Eminem.

Conclusion

As a result of this project, I believe I have generated four key insights:

1. The usage of misogynistic slurs in rap and hip hop music has been increasing.
2. The standard deviation of the usage of misogynistic slurs in rap and hip hop albums each year is greater than the standard deviation of the usage of homophobic slurs in rap and hip hop albums each year.
3. Based on the data and analysis I used, Snoop Dogg, Missy Elliott, and Eminem are the artists who use the most homophobic slurs.
4. Based on the data and analysis I used, Lil Wayne, D12, and Post Malone are the artists who use the most misogynistic slurs.
5. Based on the data and analysis I used, *Doggystyle* by Snoop Dogg, *808's and Heartbreak* by Kanye West, *Raising Hell* by Run-D.M.C. are the albums that use the most homophobic slurs.
6. Based on the data and analysis I used, *Tha Carter III* by Lil Wayne, *2001* by Dr. Dre, and *The Marshall Mathers LP* by Eminem are the albums that use the most misogynistic slurs.

References

1. <https://slate.com/culture/2009/08/the-rise-of-no-homo-and-the-changing-face-of-hip-hop-homophobia.html>
2. <https://journals.sagepub.com/doi/pdf/10.1177/1097184X08327696>