# Women in Music

It is no secret that the music industry, like many others, is dominated by men. The extent is difficult to measure and analyze due to the creative and individual nature of music. Since 2012, the USC Annenberg Inclusion Initiative has attempted to capture these inequities by analyzing inclusion on the Billboard Hot 100 Year-End Chart. In 2022, there were a total of 160 artists on the Hot 100 Year-End Chart—96.4% men, 30% women, and 0.6% gender non-binary (4). This is a 7.3 percentage point improvement since 2012 for women. Of this 30% of artists that women make up, most worked in Pop where 33% of songs were by women. Women were then most likely to work in Dance/Electronic, R&B/Soul, Country, Alternative, and least likely to work in Hip-Hop (4). The only non-binary artist on the Chart was Sam Smith who received 3 credits in Pop.

If the representation of artists seems alarming, representation in songwriting and production is even worse. Of the 451 songwriters on the Chart in 2022, 14% were women, and only 3 credits belonged to gender non-binary songwriters—all 3 again belonging to Sam Smith (12). There were a total of 231 producers credited on the Chart in 2022, 3.4% of whom were women, and only one of whom was gender non-binary (again Sam Smith) (16). The USC Annenberg report captures the inequities in the music industry, however it falls short in explaining their causes, and only addresses the most popular mainstream music. It largely ignores the way music consumption has become an individual experience through the rise of streaming.

Similarly, Italian researchers used data from Spotify, webpages, and more to assess gender discrimination among Italian singers. The study used the number of listeners on Spotify as their measure of success for each artist. They ultimately concluded there is no gender discrimination among Italian singers as there was no significant difference in number of listeners by gender, despite the average number of listeners for male singers being consistently higher than that of female singers (Nappo). Their analysis reveals a difficulty in measuring gender bias, despite the existing gender disparity in the music industry, especially using Spotify metrics. Number of listeners may not reflect success, and not all listeners are created equal—some are loyal fans while others could be generic playlist streams. Similar to the Annenberg report, the study does not reveal any forces responsible for these discrepancies, but this study also reveals the difficulty in capturing them.

While the USC Annenberg report addresses popular music, interviews from 17 “female digital musicians” describe the experience of being a self-taught musician performing online without a record label. The study focused on musicians using YouTube, Soundcloud, and Vine, however Spotify is another online platform where independent artists can promote their work. The musicians cite expectations of their physical appearance as a barrier to success. They see the differences in portrayal of male and female musicians in the music industry and rewarding of female sexualized body parts as contributing to physical appearance being a significant part of being a female musician online (Choi, 2016). This study address female musicians promoting their work online as well as artists outside of mainstream music, but focuses on individual experiences and cultivating online presence. It does address barriers to success aside from the expectation for a female musician to be beautiful, in addition to talented.

# Gender

All of the previously mentioned studies convolute sex and gender, excluding other gender identities and expressions. In her work “Queering feminist technology studies”, Landström suggests that heteronormativity influences feministic research on gender and technology. She calls for a theoretical shift in framework to account for technology’s coproduction of gender to avoid empirical research that relapses into old patterns of masculine and feminine binary.

Many studies on the classification of gender or technology and gender fall into this archaic thinking. For example, a study from Verizon, Yahoo, and Worcester Polytechnic Institute attempted to create a machine learning model to predict a user’s gender based on their username. They only allowed for two outcomes, though: male or female. They found a character-based logistic regression approach was most accurate for gender classification, though there are problems with all attempted gender classification models (Hu). Their work demonstrates name is not enough information to make accurate gender predictions as well as the shortcomings of binary gender classification.

One such example is a research study from Verizon, Yahoo, and Worcester Polytechnic Institute that attempted to create a machine learning model to predict a user’s gender based on their username, while only allowing for two outcomes: male or female (Hu). The research was motivated by a desire to determine a user’s gender when it was not supplied in account creation in order to use this information as input data for recommendation algorithms. They found a character-based logistic regression approach was most accurate for gender classification, though there are problems with all attempted gender classification models (Hu). Their work demonstrates name is not enough information to make accurate gender predictions as well as the shortcomings of binary gender classification.

* Add in Rena’s “Baking Gender Into Social Media Design” about how gender is no longer a required attribute in account creation but largely controls user experience

Classifying gender into a binary is ignorant of other identities, but classifying gender at all can be misleading.

* Societal construct of gender could start here and then flow into this next paragraph

# Gender and Algorithmic Bias

This ignorance breeds bias in our algorithms which perpetuates the heteronormativity, ignorance, and oppression which contribute to it. Algorithmic bias extends far beyond the context and application of the algorithm itself. As technology becomes more ingrained in society, it becomes another part of society which produces gender (Landstrom).

* And then examples from the gender data gap book about algorithmic bias :

This algorithmic bias is already apparent in daily life. Criado Perez’s “Invisible Women” details how the gender data gap, the systemic and historic lack of data collected on or about women, perpetuates the oppression that created it. Voice recognition in cars, for example, relies on an algorithm “trained on large databases of voice recordings…dominated by…male voices” as such there are “70% more likely to accurately recognize male speech than female speech” (162). She warns that algorithms trained on “gap-ridden corpora”, or input data, leaves the impression that men are the default, and that these machines are not only reflecting but also amplifying our biases (164).

These algorithms are already distorting our perceptions about reality and reinforcing patriarchal ideas. Recommendation algorithms have only amplified these effects—their bias generating biased output contributing to the initial bias. YouTube’s recommendation algorithm has already been seen to favor feminized content, creating a vicious cycle that rewards oppressive norms. YouTube systematically promoted feminized material by rewarding and promoting feminized content in beauty vlog recommendations. [add more here] When failing to account for potential bias, developers create algorithms with unintended side effects which amplify existing societal problems such as discrimination, gender bias, and gendered societal expectations (Bishop).

These unintended effects have unintended effects themselves. Such as machine learning gender bias producing more accurate results for male authors than female authors which only perpetuates the gendered data gap and feeds into the inaccurate or inadequate results for women (Thelwall). They also have psychological effects. Researchers at Stanford found women receiving stereotypically “feminine” career recommendations had lower estimates of leadership ability than those who received traditionally “masculine” career recommendations and believed the recommendations were based on internal characteristics (French). An understanding of these systems, algorithms with implicit bias, moderates its effect on sense of self but cannot remove them entirely (French).

These algorithms are being fed biased data as a result of gender-based oppression which is generating biased output only contributing to the problem.

# Spotify

Spotify algorithms have been analyzed in a variety of ways in the past, though little is known about the gender discrepancy in their output. Researchers at Spotify and University of Toronto found that algorithmic recommendations are more effective for users with lower diversity and are associated with reduced diversity of consumption. Diversity of consumption, however, is important for long-term user retention (Anderson). Spotify’s recommendation algorithm is constantly evolving to address this nuanced tension between short-term enjoyment and accurate recommendations with diverse listening and long-term user retention (Anderson). Diversity, here, however, applies to the diversity of the music being listened to, not of the performer, songwriter, or producer. More diverse consumption would mean consumption across a wide variety of genres, not necessarily a wide variety of artist identities.

* Transition about listening to more diverse artists to the importance of representation

Gender is hypothesized to be part of Spotify’s “black box” algorithm as demographic

In the context of podcasts, researchers found that personalized recommendations increased podcast streams but decreased individual listening diversity (Holtz). Another group which received recommendations based on popular podcasts among users in their demographic group …

Gender bias in Spotify’s “black box” algorithm has also been analyzed.

* Some background about Spotify and black box algorithms generally
* More studies about Spotify recommendations and gender
* The end

Researchers at Linkoping University attempted to understand user’s gender’s impact on Spotify’s “black box” algorithm. They performed 288 bot experiments within four genres—rock, gospel, rnb/hip-hop, and dance-electronic—with each bot registering as a male or female user and listening to only the top 10 songs out of the top 100 songs of their given genre. The majority of bot users were given the same recommended artists regardless of their account’s gender, though gender-skewed recommendations were found in rock as well as rnb/hip-hop and dance/electronic (Friksson). This experiment concludes that account gender is not a heavily weighted factor in Spotify’s recommendation algorithm but does not further analyze these “gender-skewed” recommendations by category.

Another study attempted to understand how Spotify recommendation algorithms present gender while also presenting music by analyzing the listening experience of 80 young adults in Moscow and Stockholm. The visual user interface as well as recommended song information were analyzed for three of Spotify’s common recommendations—Related Artists, Discover, and Browse. Related Artist recommendations are artists recommended based on a current artist the user is viewing. These recommendations were found to often have artists of the same gender and demographic group as the original. The Discover playlist is a playlist of 30 songs generated weekly for each user individually and was less demographically cohesive, though still recommended music similar in genre, gender, and race. Browse algorithms are general playlists available for large communities of users and were the most diverse. The study confirms that the Discover playlist is mostly unique to each user but did not address what other factors might determine the recommendations made there.

Many attempts have been at “reverse engineering” Spotify’s recommendation algorithms.

It is suggested that Spotify uses content-based and collaborative filtering to generate track representations. Content-based filtering meaning the content of each track, such as artist name, songwriting credits, and genre, but also includes metrics Spotify calculates based on these attributes such as danceability as well as a semantic meaning estimated with natural language processing models. Collaborative filtering describes the track’s connection to other tracks and what other users with similar tastes are listening to. Each user on Spotify also has their own “taste profile” or Spotify’s understanding of what the user likes based on their listening history. While Spotify reveals a lot to users and artists about how its algorithms work, it is currently unknown whether artist gender is meaningfully included in the content-based filtering, and it is also unknown whether there is a data gap in music from female-identifying musicians based on the near infinite amount of content on Spotify. As such, it is difficult to know if there is gender bias in Spotify’s recommendation algorithms and where it is coming from.

<https://www.music-tomorrow.com/blog/how-spotify-recommendation-system-works-a-complete-guide-2022>

Given the gendered data gap as well as unavoidable gender bias, I wish to analyze Spotify’s recommendation algorithm for Discover Weekly, the most personalized of Spotify’s recommendation algorithms, for gender bias. Previous studies have shown that demographic information is not as heavily weighted in the Spotify algorithms as listening history. As such, I will set up an equal amount of male and female profiles to confirm this finding. Spotify’s algorithm is difficult to analyze for bias as its inputs and their weights are largely unknown. As such, I will use a variety of accounts with a variety of identities and musical preferences to identify larger trends in the algorithms as a whole.

* Understand where this gender bias may be more or less present