Molly Siebecker

DATA 698

Prof. Hagstrom

May 20th, 2025

Gender Identity and Emotional Expression in Popular Music Lyrics

I. Abstract

Popular music has long been used as a barometer of American culture and society, reflecting the national mood. Prior research that utilizes sentiment analysis on lyrics has focused on genre and the gender of the artist as explanatory variables. In this study, the gender of the songwriter is taken into account to determine if there is an association between the gender identity of a songwriter, or musical artist, and the emotional content of the lyrics, as determined by sentiment analysis methods. A sentiment classifier was used to determine confidence in the expression of six major emotions in the top 100-charted songs of each of the last 60 years. The rate of emotion-word usage was also calculated. It was found that there are associations between the emotional content of songs and both the gender of the artist and the songwriter, particularly involving the expression of anger, where female artists and female songwriters are under-represented. Generally, the gender of the artist has more of an association with emotional content than the gender of the songwriter. The songs of female artists also display a greater rate of emotion-word usage, and the classifier predicts greater confidence in the emotional content of their songs.

II. Introduction

While popular music has historically been dominated by male artists, a 2022 industry study on inclusion metrics found that in recent years, female artists have made gains in representation, in terms of chart rankings and Grammy awards (Smith et al, 2023). There is a growing cultural narrative that women dominate the current state of pop music. In the main telecast of the Grammys, which includes both the general field and specifically-highlighted genres (generally considered to be the most popular or hold the most weight), women and non-binary artists accounted for between 70% and 100% of the award recipients over the past two years (Recording Academy, 2024-2025). Cultural commentators, industry professionals, and psychologists have opined that the increase in popularity and cultural significance can be attributed in part to female consumers, and that those female consumers are drawn in by lyrical content and emotions (Butler, personal communication, 2024; Gold, personal communication, 2023).

However, music production still displays significant gender disparity, particularly for female songwriters, as only 12.8% of songwriters are women (Smith et al, 2023). Top male songwriters have approximately twice as many credits as the top female songwriters have (Ibid). The ten male songwriters with the most credits "are responsible for 23.9% of the 1,100 most popular songs from 2012 to 2022" (Ibid). Using natural language processing techniques to determine if there is an association between the gender identity of the songwriter or musical artist and linguistic features of lyrics could potentially have implications for industry practices and gender representation in songwriting.

III. Literature Review

Researchers have noted a shift in the sentiments expressed in the lyrics of popular music over the past half-century. Positive sentiments have declined, and negative sentiments have increased (Napier and Shamir, 2018). In that study, IBM Watson's Tone Analyzer was used to assign specific sentiments to lyrics, which found that individual negative sentiments like anger and sadness had generally increased and positive sentiments like joy had generally decreased (2018). Studies have similarly found that song lyrics have become more self-focused, more socially disconnected, more angry and antisocial, and less positive (Dewall et al, 2011).

Musical genre has its own associations with gender and with emotion. The previously mentioned study discussed associations between linguistic features and genre, and found that pop music showed the greatest association with positive affect, rock music with the use of the 1st person plural, and hip-hop/R&B with the use of the 1st person singular and antisocial indicators (Dewall et al, 2011). Some genres, like rock, are perceived to be more "masculine," and women are under-represented as band members and instrumentalists (Schaap and Berkers, 2014), whereas women are more expected to be vocalists and pop music artists.

Women and men display key differences in the expression of emotion. Though women self-report experiencing anger just as much as or more than men, men "express more anger through vocal, facial, or behavioral modalities," and observers are more able to correctly identify men's facial expressions of anger than women's (Brody and Hall, 2000). They also found that women report experiencing greater emotional intensity and are more likely to communicate

about their emotions. Surveys show that men report more feelings of calm and excitement, and women report more feelings of anxiety and sadness (Simon and Nath, 2004). These differences are generally found to be the product of socialization and not of innate biological differences (Brody and Hall, 2000).

Studies have also long established that men and women write differently in everyday communication, as women use more emotion words and men use more object references, among other lexical differences (Schwartz et al, 2013). Women also "use more words related to psychological and social properties" (Newman et al, 2008). These differences are found to extend beyond everyday communication and into formal written texts (including both fiction and nonfiction) (Argamon et al, 2003). Machine learning techniques have been used to categorize the gender of the author of a formal written text with approximately 80% accuracy (Koppel, 2002).

IV. Methods

Data was collected by scraping the titles and artists of the Billboard Year-End Top 100 from 1965-2024 from Wikipedia[1], and then using API requests to obtain the corresponding lyrics from Genius and the songwriter(s) from MusicBrainz. MusicBrainz is an open-source database of music metadata that contains songwriter information under three categories: composer, songwriter, and lyricist. The variable "songwriter" was created by combining both songwriter and lyricist information, and the variable "lyricist" was created with lyricist information only. Very few songs had lyricist information, and this variable was ultimately discarded in analysis. The genre of each song was obtained as an additional variable by scraping from Wikipedia.

[1] Code for scraping the songs from Wikipedia was modified from Kaylin Pavlik, "50 Years of Pop Music."

The gender identities of both artists and songwriters were scraped from first Wikipedia and then MusicBrainz, for those that could not be found on Wikipedia. Both Wikipedia and MusicBrainz use gender identity self-identification that recognizes non-binary identity. A function used logical reasoning to check keywords to see if the artist was in fact a musical *group*, in which case "group" was assigned for identity. Then, the function checked for key words to see if they were non-binary. If neither "group" nor "non-binary" applied, the function used the first pronoun present in the article to assign gender identity. Both artist and songwriter gender identity were stored in separate dictionaries, which were later mapped to the songs in the data frame.

Some missing data or errors were fixed manually; for example, the Genius API returned some translations of lyrics that were originally written in English, and these were replaced with the English lyrics manually. Instrumental songs were identified and labeled manually. Google's Compact Language Detector 3 was used in R to detect and label the language of each song. A select few songs that were written using multiple languages or dialects (for example, "Shy Guy" by Diana King) were also labeled manually. This led to four new variables: instrumental, detected_language, is_non_english, and mixed_language. Instrumental songs and songs fully or partially written in a language other than English were excluded from analysis.

The resulting data frame, after mapping genders, contains over 6,000 songs[2] for which key variables include the year charted, rank, song title, artist(s) and their gender(s), genre, lyrics of the song as raw text, and songwriter(s) and their gender(s). Additional variables were created to

[2] Some charting songs from earlier decades were Side A/ Side B singles, which were separated into two separate songs to preserve the distinct lyrical content.

count the number of songwriters on each song, and the count and percentage of songwriters by gender. Only 46 songs had at least one non-binary songwriter credited, which, as less than 1% of the total sample, was considered too small a subsample for meaningful inferences to be drawn, due to the extreme imbalance of the classes. An additional variable, artist_songwriter_combo, was created which defined the combination of genders in songs that had exactly one artist and exactly one songwriter.

The raw text of the lyrics was preprocessed by standardizing whitespace, to prepare the data for natural language processing. A separate variable was created which stored the tokens of each set of lyrics in a list. The DistilBERT-base-uncased-emotion classification model from Hugging Face was applied for sentiment analysis[3] (Savani, 2020). The model was applied to the lyrics as text, which did not require stop words to be removed. Six new variables were created, each one a confidence score, for the emotions anger, sadness, joy, love, fear, and surprise. The maximum confidence score in each row was used to create the variables primary_emotion and primary_emotion_score.

In addition to the emotional content of the lyrics, the lyrical expression was evaluated by finding the rate of emotion-laden words per song. The NRC Word-Emotion Association Lexicon was used to determine the number of emotion-laden words in each song by counting the number of tokens in each song that appear in the lexicon. This was used in conjunction with the total number of tokens per song to determine the rate of emotion-laden words per 100 words.

[3] The text was truncated to the first 512 tokens to fit model requirements. This was thought not to substantially affect overall results, as popular music songs typically maintain one emotion throughout.

Potential associations between the categorical variables of songwriter gender, artist gender, artist/ songwriter combo and primary emotion were assessed using chi-square tests. There were no songs which were written by one non-binary songwriter, so this level was not present in the relevant contingency table of songwriter gender and primary emotion. In the contingency table for artist gender and primary emotion, multiple cells in the row for non-binary gender had a count less than 5, violating the assumptions of the chi-square test, so this level was excluded. In the contingency table for artist/ songwriter combo and primary emotion, non-binary gender, fear, and surprise violated the ≥ 5 assumption and were excluded.

Potential associations between the categorical variables of gender and the quantitative variables of rate of emotion-laden words, primary emotion score, and the six emotion-specific scores were assessed using Mann-Whitney U tests. T-tests were considered and rejected because the assumption of normality was not met in any of the 8 quantitative variables.

V. Results

The results of the three chi-square tests are summarized in Table 1 on the following page. There was a statistically significant association present between artist gender and primary emotion, and an association at the threshold of statistical significance between songwriter gender and primary emotion. Full contingency tables are included in the Appendix.

Table 1. Results of Chi-Square Tests on Gender and Primary Emotion

Variables	χ^2 (df)	*p*
Songwriter Gender x Primary Emotion	11.06 (5)	0.05*
Artist Gender x Primary Emotion	12.47 (5)	0.03**
Songwriter/Artist Combo x Primary Emotion	9.73 (6)	0.14

The standardized residuals between the actual and expected frequencies are displayed in the figures below. Only the standardized residual of 2.87 for women songwriters and surprise met the threshold of $\pm$ 2 for statistically significant over- or under-representation. For artist gender and primary emotion, the trends visible throughout the standardized residuals are more diffuse.

Figure 1: Associations between Songwriter Gender and Primary Emotion

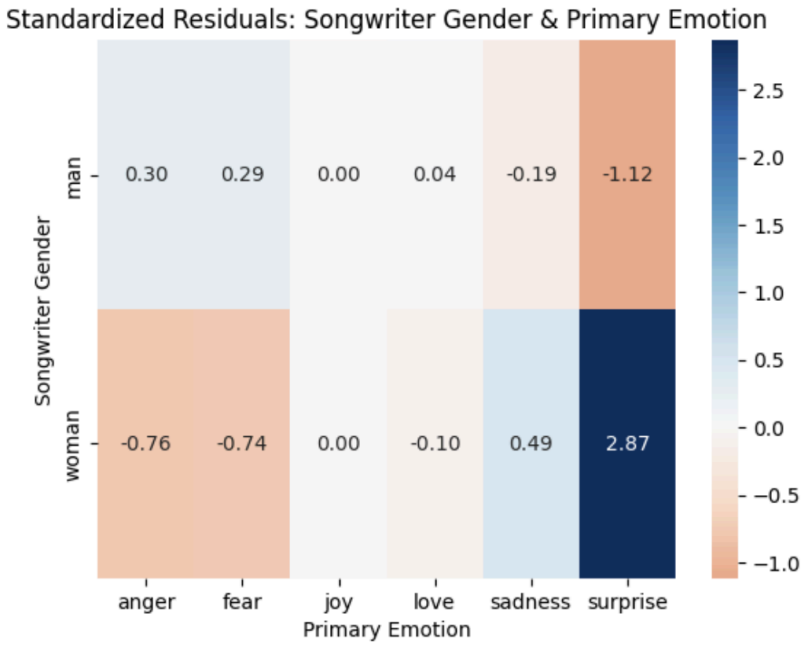
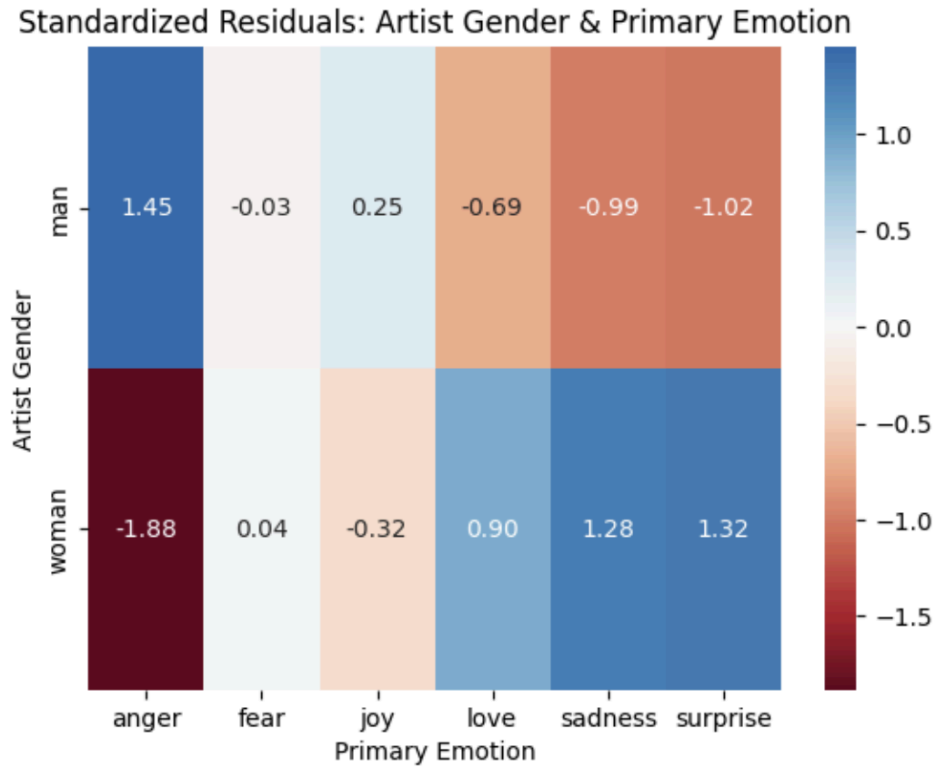


Figure 2: Associations between Artist Gender and Primary Emotion



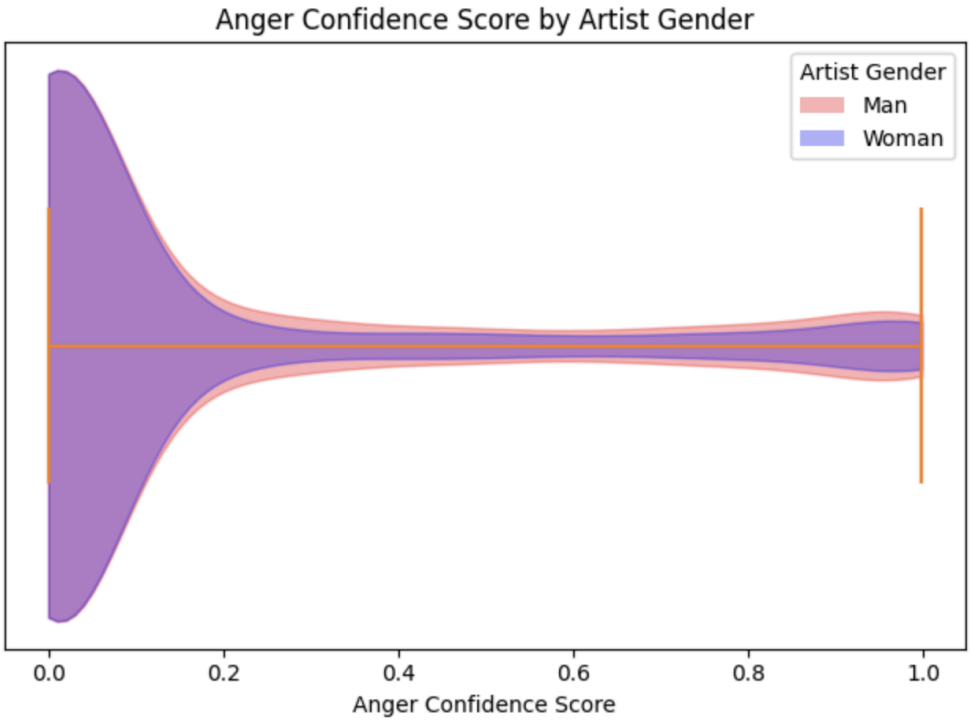
Out of the 16 Mann-Whitney U tests that were performed, 5 met or approached ($p \approx 0.058$) the threshold for statistical significance. All involved comparisons on artist gender and not songwriter gender. These associations are between artist gender and emotion-laden word rate, primary emotion confidence score, and the confidence scores for anger, fear, and surprise; these selected results are summarized in Table 2, and full results are available in the Appendix.

Table 2. Selected Results of Mann-Whitney U Tests Comparing Male and Female Artists

Variable	Male Artists Median (Mean ± SD)	Female Artists Median (Mean ± SD)	*U*	*p*
Emotion-laden word rate (per 100 words)	9.350 (9.997 ±4.510)	9.708 (10.414 ±4.800)	768,974.5	0.058*
Primary Emotion Confidence Score	0.955 (0.863 ±0.172)	0.974 (0.878 ±0.166)	752,973.0	0.006**
Anger	0.014 (0.194 ±0.316)	0.008 (0.150 ±0.292)	879,195.0	<0.001***
Fear	0.002 (0.056 ±0.189)	0.002 (0.057 ±0.194)	854,383.0	0.008**
Surprise	0.001 (0.009 ±0.070)	0.001 (0.011 ±0.081)	848,844.5	0.018**

Figure 3 on the following page shows the relative distributions of anger confidence score by artist gender, which showed the greatest association between gender and any quantitative variable. Figures showing the relative distributions of the other statistically significant associations are included in the Appendix.

Figure 3: Violin Plot of Confidence Score in Anger, Disaggregated by Artist Gender

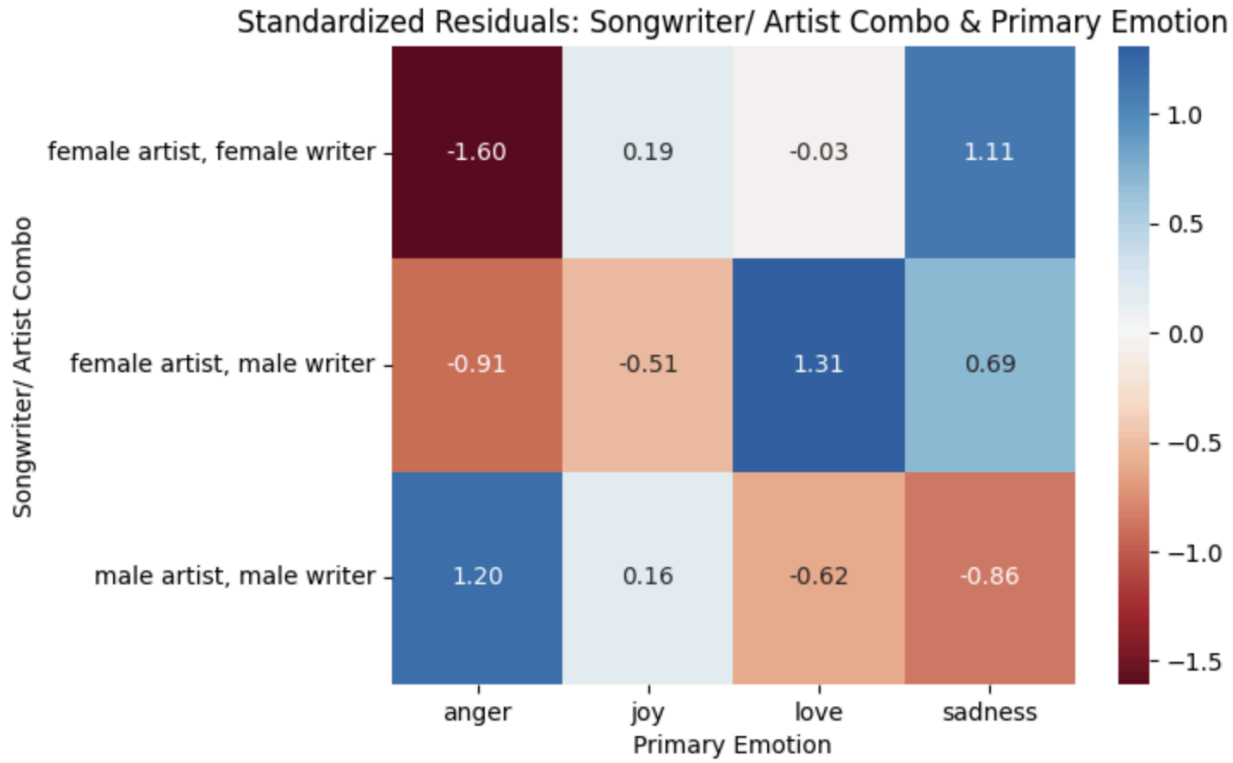


VI. Discussion

The association between artist gender and primary emotion is statistically significant, but in individual gender/ emotion pairings, this association is relatively diffuse, as none of the standardized residuals were beyond $\pm$ 2. However, several categories still showed meaningful trends when broken down by artist gender. Male artists are over-represented in songs displaying anger, and female artists are over-represented in songs expressing sadness and surprise. Though technically not statistically significant, this is in line with previously discussed prior research which has found that men express more anger, and women report more feelings of sadness.

Though the association between artist/songwriter combination and primary emotion did not reach the threshold for statistical significance, again, there are interesting and potentially meaningful patterns in the data.

Figure 4: Associations Between Songwriter/Artist Combination and Primary Emotion

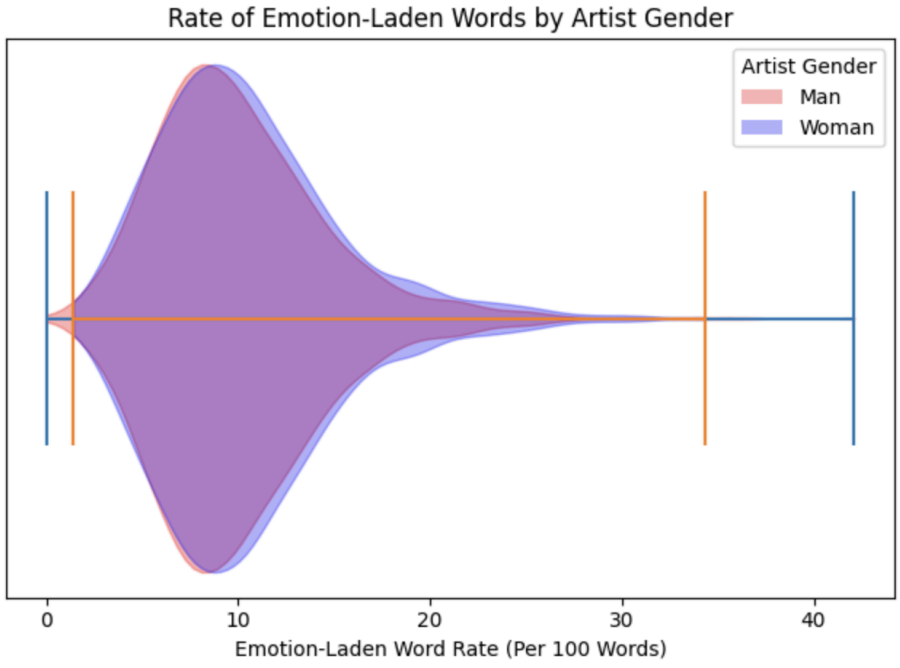


The category with the strongest level of over- or under-representation is songs expressing anger by female artists with female writers, even more under-represented than songs expressing anger by female artists with male writers, which suggests that this association is driven less by stereotypes held by the writer about the gender of the artist, and more by the identity and expression of the writer. Songs by male artists with male writers over-represented in anger, consistent with previous findings and prior research around men's expression of anger. Likewise, sadness is over-represented in songs by female artists with female writers, and under-represented

in songs by male artists and male writers, consistent with research about women's greater self-reporting of sadness and about traditional expectations of masculinity and the repression of sadness. Interestingly, the greatest level of over-representation is love songs, in the cross-gender category of female artists with male writers.

The association between the gender of the artist and the rate of emotion-laden words is supported by previous findings regarding the stylistic differences between men's and women's writing. Figure 5 below illustrates how songs by female artists use emotion-words at a greater rate. However, unlike previous research about men's and women's writing, this association is present with the gender of the *artist*, not of the songwriter. This suggests that this stylometric feature is not a fixed characteristic of the writing of an author, but is related to gender as a performative act, as described in feminist theoretical frameworks (Butler, 1988).

Figure 5. Distribution of Emotion-Laden Word Rate by Artist Gender



There were a few challenges with this study; first are the potential limitations of the sentiment classifier. It does have high accuracy rates in labeling emotions, but has not been trained on figurative language in the ways that some LLMs have been (Savani, 2020). Additionally, it is possible that it would show bias in text written in African American Vernacular English, which could skew the results for songs that utilize AAVE (Mehrabi, 2021). Another limitation is that this research does not take into account the impact of genre, which may have its own associations with emotion or with artist or writer gender. Genre information for each song was collected, but these classifications were excessively granular and frequently overlapping, like "country pop" and "pop rock," in ways that would hinder creating independent genre categories for statistical analysis. Finally, the subset of songs that only had one writer skewed towards earlier decades because the number of songwriters per song has been increasing over the years.

VII. Conclusion

There are statistically significant associations between both artist gender and songwriter gender and the primary emotion of a song. Male artists, but not male songwriters, tend to be over-represented in songs expressing anger. Both female artists and female songwriters are over-represented in songs expressing surprise, and female artists only are over-represented in songs expressing sadness.Though both artist gender and songwriter gender displayed associations with the primary emotion of the song, only artist gender displayed any significant association with the emotion-word rate and individual emotion confidence scores, which suggests that the gender of the artist is a potentially more significant explanatory variable than the gender of the songwriter, perhaps related to the performative aspects of gender. Finally, anger

was the emotion that consistently showed the greatest statistically significant association with artist gender, as seen in both the chi-square and Mann-Whitney U tests.

In future research, the interactions between both genre/gender and genre/emotion should be investigated. Future studies would also benefit from repeating this type of analysis with a larger data set of songs that feature one artist and/or one writer, not only songs that necessarily made it into the Top 100 of the year, which, particularly in recent years, frequently feature multiple artists and multiple writers. Finally, it would be valuable to determine if there is an association between whether or not the artist and the songwriter are the same person, and the emotional expression of the song.

VIII. References

Argamon, Shlomo, Jonathan Fine, and Anat Rachel Shimoni. "Gender, Genre, and Writing Style in Formal Written Texts." *Text*, vol. 23, no. 3, 2003, pp. 321–346. https://doi.org/10.1515/text.2003.014.

Brody, Leslie R., and Judith A. Hall. "Gender and Emotion in Context." *Handbook of Emotions*, edited by Michael Lewis, Jeannette M. Haviland-Jones, and Lisa Feldman Barrett, 2nd ed., Guilford Press, 2000, pp. 395–408. https://coachingadistancia.com/wp-content/uploads/2020/09/61-262-Handbook-of-emotions-Michael-Lewis-Jeannette-M-Haviland-Jones-And-Lisa-Feldman-Barrett.pdf.

Butler, Judith. "Performative Acts and Gender Constitution: An Essay in Phenomenology and Feminist Theory." *Theatre Journal*, vol. 40, no. 4, 1988, pp. 519–531. https://doi.org/10.2307/3207893.

Butler, Tonya. Interview by Christina Pazzanese. "It's a Man's World? Definitely Not This Year." *The Harvard Gazette*, 30 Jan. 2024, https://news.harvard.edu/gazette/story/2024/01/watershed-moment-for-women-in-music business-grammys/.

Compact Language Detector 3. *Google*, version 1.6.1, 2025, https://cran.r-project.org/web/packages/cld3/index.html.

DeWall, C. Nathan, et al. "Tuning In to Psychological Change: Linguistic Markers of Psychological Traits and Emotions Over Time in Popular U.S. Song Lyrics." *Psychology*

of Aesthetics, Creativity, and the Arts, vol. 5, no. 3, 2011, pp. 200–207. https://doi.org/10.1037/a0023195.

Genius API. *Genius*, 2025, https://docs.genius.com/.

Gold, Alexandra, et al. Interview by Christina Pazzanese. "So What Exactly Makes Taylor Swift So Great?" *The Harvard Gazette*, 2 Aug. 2023, https://news.harvard.edu/gazette/story/2023/08/so-what-exactly-makes-taylor-swift-so-great/.

Koppel, Moshe. "Automatically Categorizing Written Texts by Author Gender." *Literary and Linguistic Computing*, vol. 17, no. 4, 2002, pp. 401–412. https://doi.org/10.1093/llc/17.4.401.

Mehrabi, Ninareh, et al. "A Survey on Bias and Fairness in Machine Learning." *ACM Computing Surveys (CSUR)*, vol. 54, no. 6, 2021, pp. 1–35. https://arxiv.org/abs/1908.09635.

Mohammad, Saif, and Peter Turney. *NRC Word-Emotion Association Lexicon*, version 0.92, 2011, https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm.

MusicBrainz API. *MusicBrainz*, 2025, https://musicbrainz.org/.

Napier, Kathleen, and Lior Shamir. "Quantitative Sentiment Analysis of Lyrics in Popular Music." *Journal of Popular Music Studies*, vol. 30, no. 4, 2018, pp. 161–176. https://doi.org/10.1525/jpms.2018.300411.

Newman, M. L., et al. “Gender Differences in Language Use: An Analysis of 14,000 Text Samples.” *Discourse Processes*, vol. 45, no. 3, 2008, pp. 211–236. https://doi.org/10.1080/01638530802073712.

Pavlik, Kaylin. “50 Years of Pop Music.” 2016, https://www.kaylinpavlik.com/50-years-of-pop-music/.

Savani, Bhadresh. *distilbert-base-uncased-emotion*. Hugging Face, 2020, https://huggingface.co/bhadresh-savani/distilbert-base-uncased-emotion.

Schaap, Julian, and Pauwke Berkers. “Grunting Alone? Online Gender Inequality in Extreme Metal Music.” *IASPM Journal*, vol. 4, no. 1, 2014, pp. 101–102. https://doi.org/10.5429/675.

Schwartz, H. Andrew, et al. “Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach.” *PLOS ONE*, vol. 8, no. 9, 2013, e73791. https://doi.org/10.1371/journal.pone.0073791.

Simon, Robin W., and Leda E. Nath. “Gender and Emotion in the United States: Do Men and Women Differ in Self‐Reports of Feelings and Expressive Behavior?” *American Journal of Sociology*, vol. 109, no. 5, 2004, pp. 1137–176. https://doi.org/10.1086/382111.

Smith, Stacy, et al. “Inclusion in the Recording Studio? Gender & Race/Ethnicity of Artists, Songwriters, & Producers across 1,100 Popular Songs from 2012 to 2022.” *USC Annenberg Inclusion Initiative*, 2023,

https://assets.uscannenberg.org/docs/aii-inclusion-recording-studio-jan2023.pdf.

Wikipedia contributors. Wikipedia. The Wikimedia Foundation, 2025, https://www.wikipedia.org/.

VIX. Appendix

Figure A1. Number of Songwriters Over Time

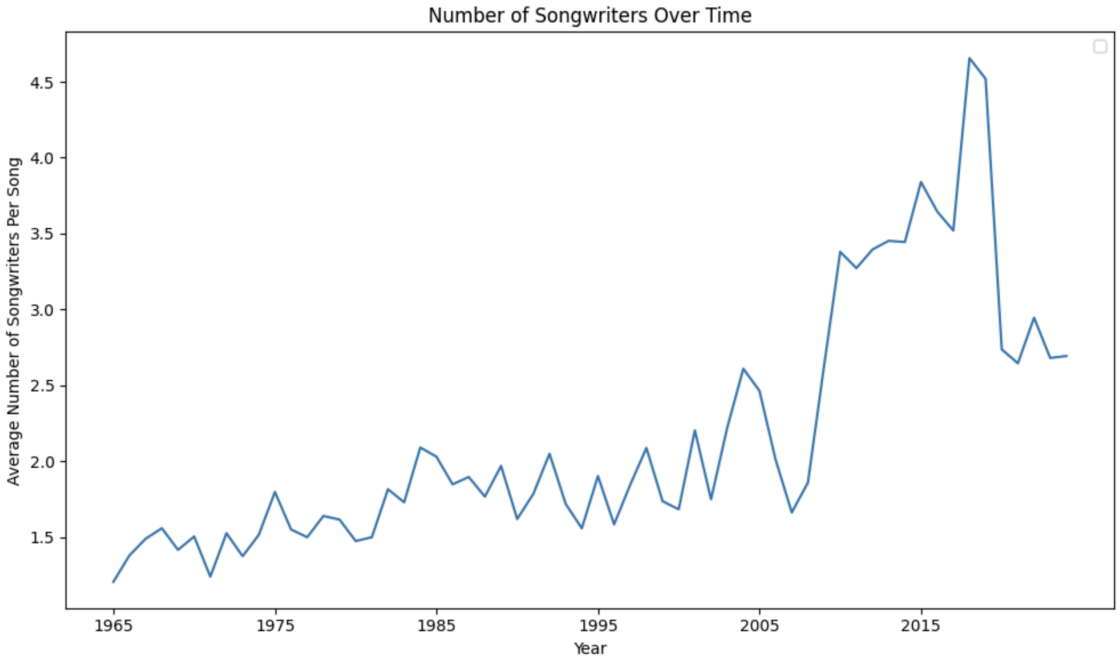


Figure A2. Composition of Songwriting Teams

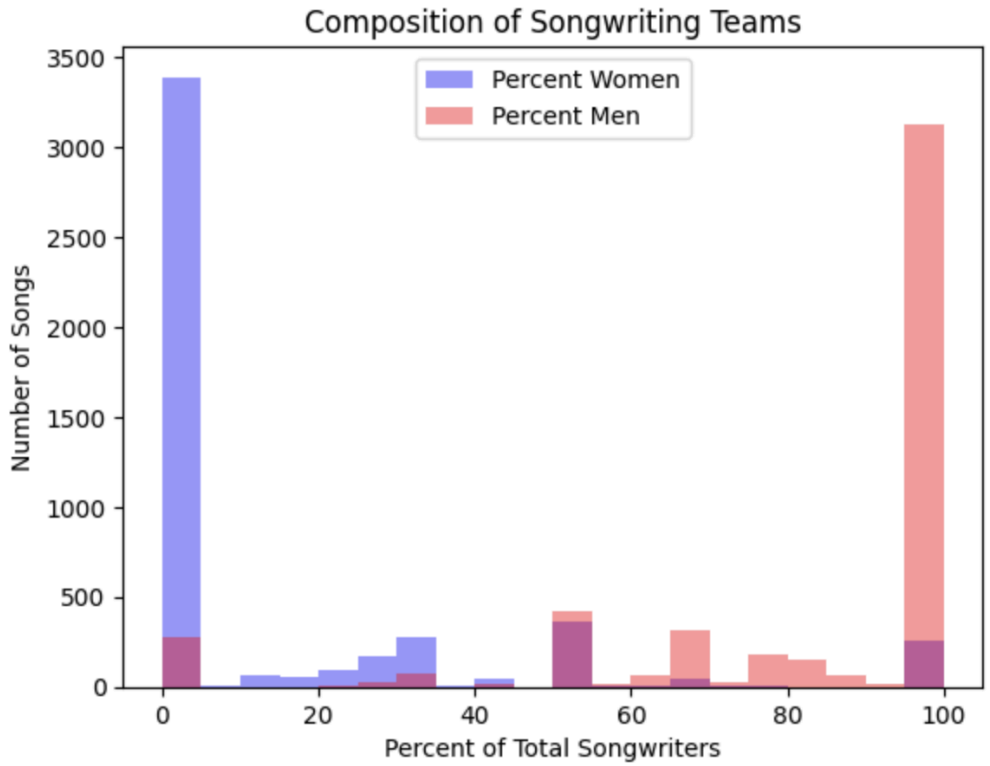


Figure A3. Distribution of Emotion-Laden Word Rate

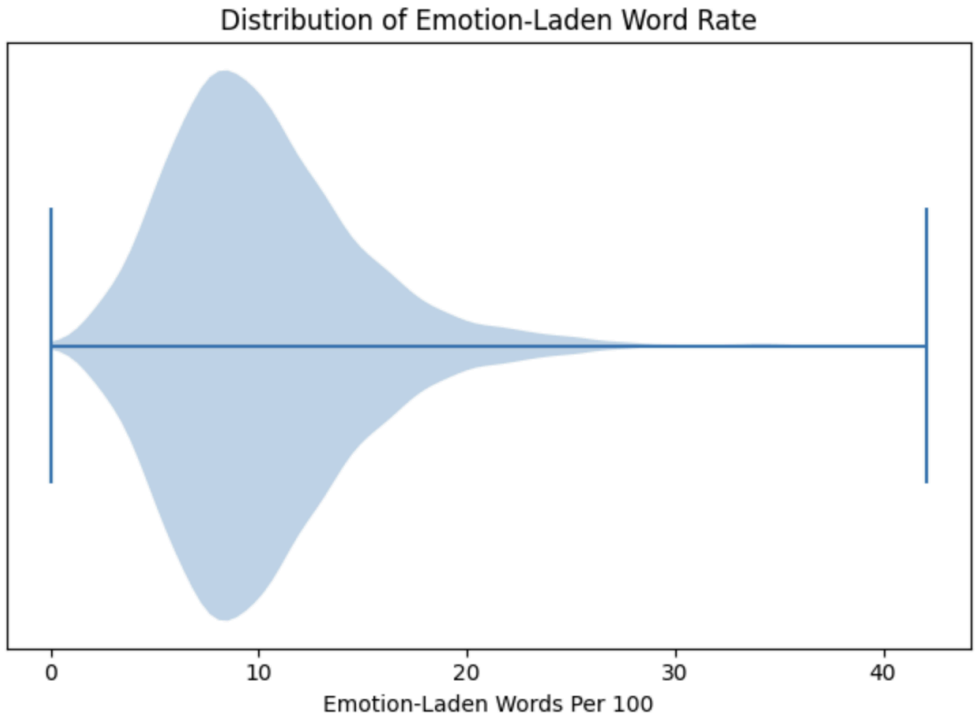


Figure A4. Distribution of Primary Emotion Confidence Score

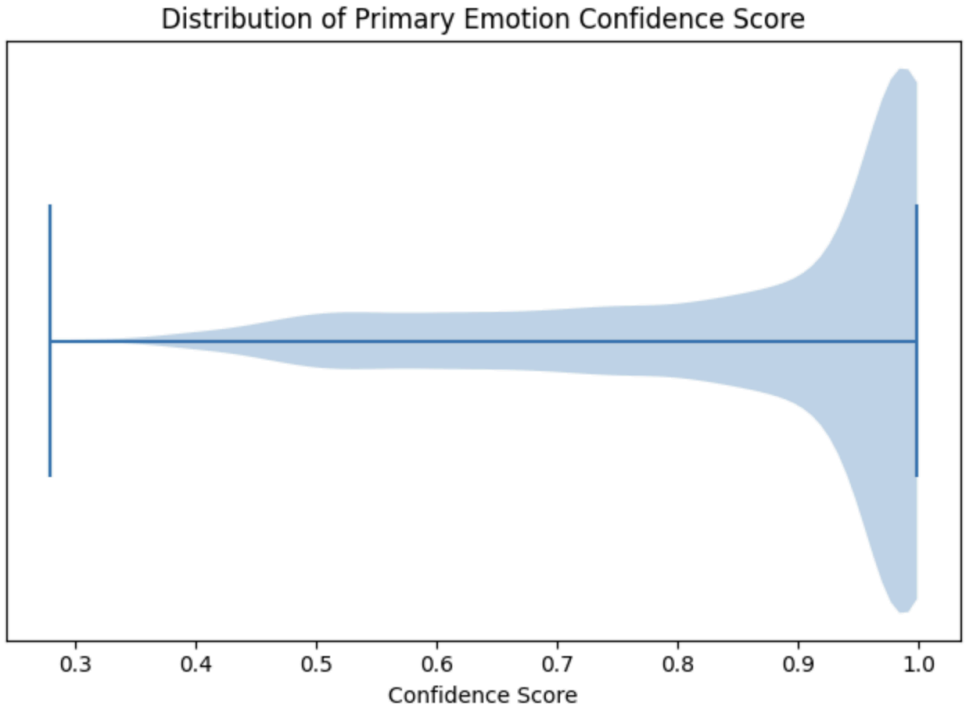


Figure A5. Distributions of Confidence Scores By Emotion

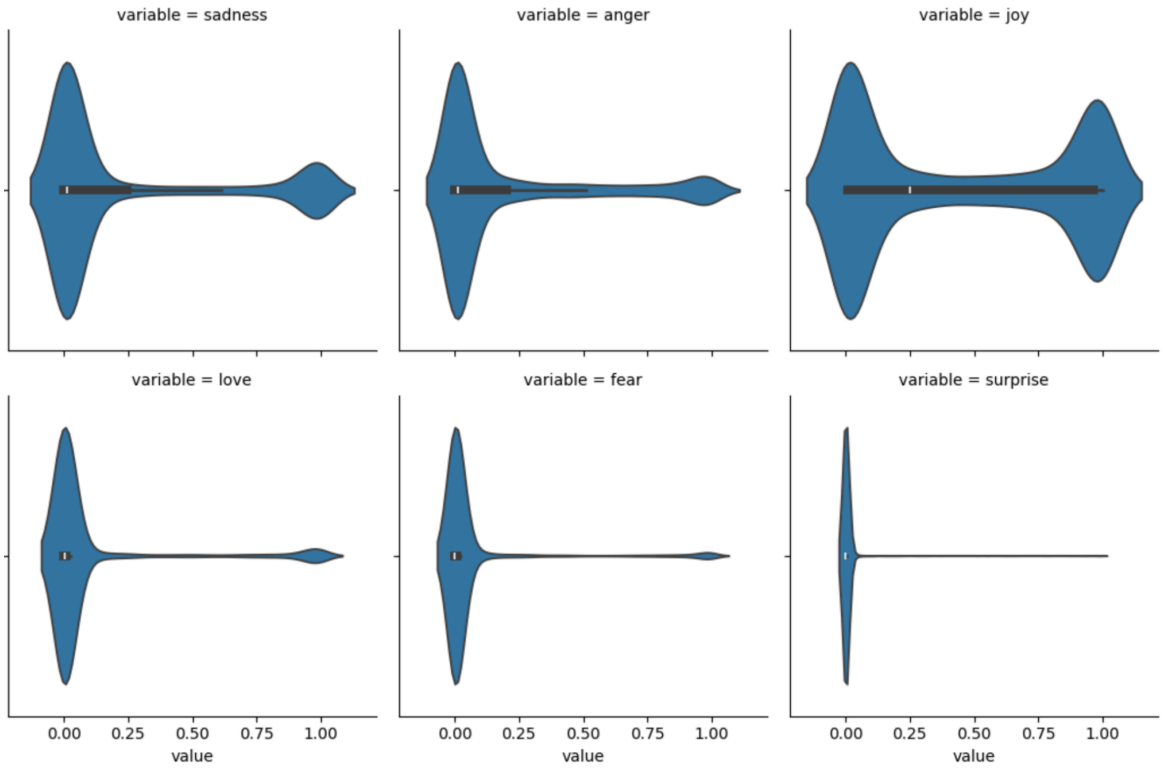


Figure A6. Distribution of Primary Emotion Confidence Score by Artist Gender

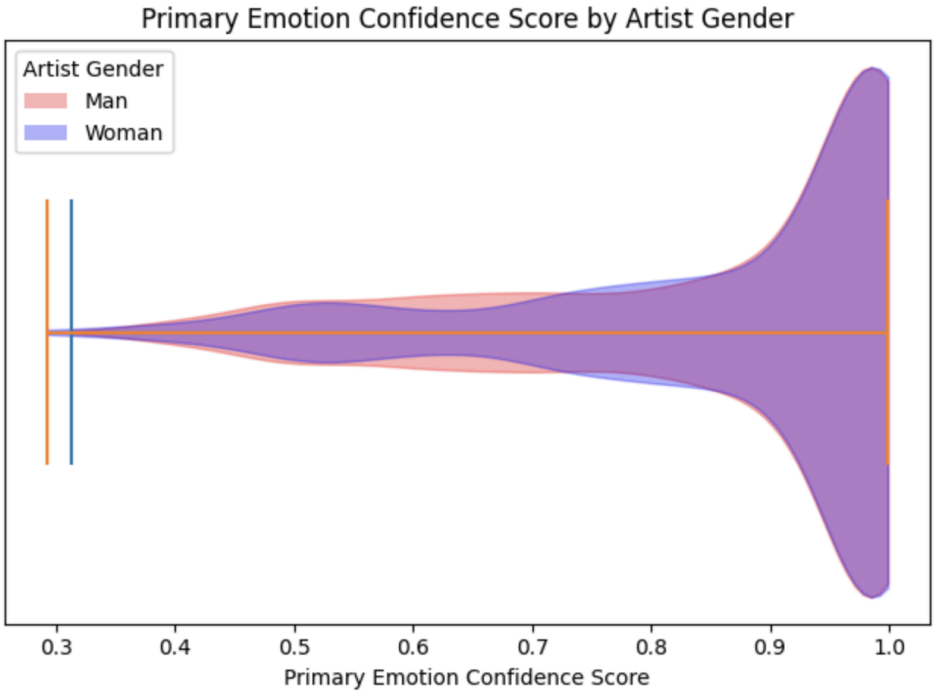


Figure A7. Distribution of Fear Confidence Score by Artist Gender

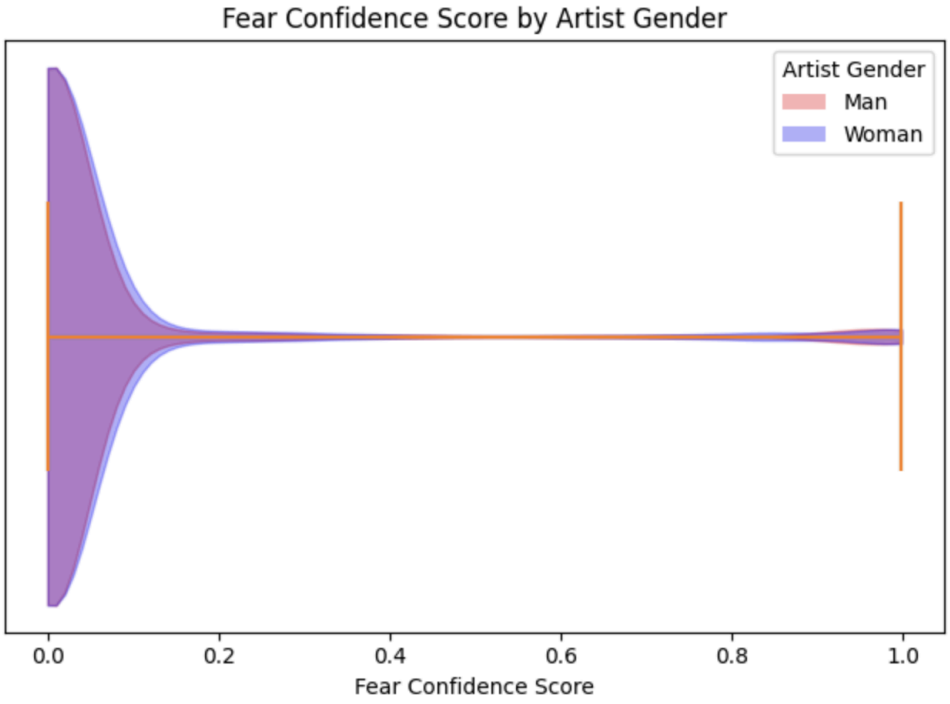


Figure A8. Distribution of Surprise Confidence Score by Artist Gender

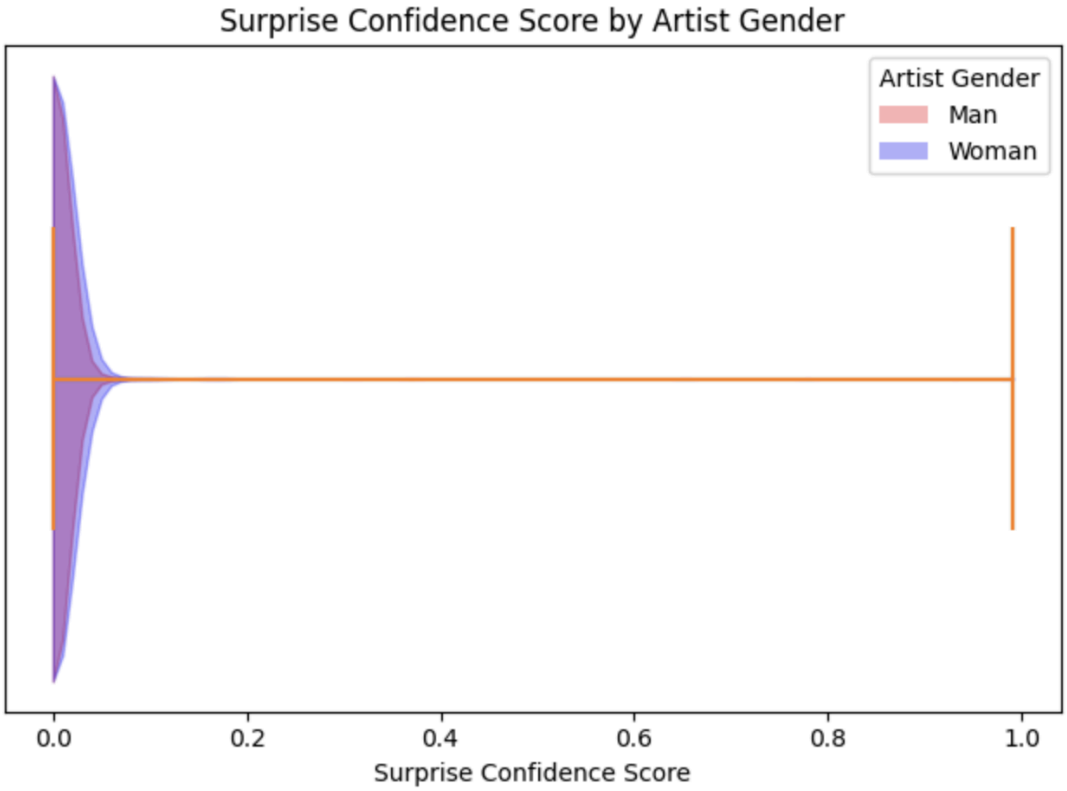


Table A1. Contingency Table: Songwriter Gender and Primary Emotion

		Primary Emotion					
		Anger	Fear	Joy	Love	Sadness	Surprise
Songwriter Gender	Man	225	86	594	122	317	9
	Woman	29	10	90	18	52	6

Table A2. Contingency Table: Artist Gender and Primary Emotion

		Primary Emotion					
		Anger	Fear	Joy	Love	Sadness	Surprise
Artist Gender	Man	225	86	594	122	317	9
	Woman	29	10	90	18	52	6
	Non-binary	5	6	7	0	8	0

Table A3. Contingency Table: Songwriter/Artist Combo and Primary Emotion

		Primary Emotion					
		Anger	Fear	Joy	Love	Sadness	Surprise
Songwriter / Artist Combo	Female artist, female writer	14	4	60	11	36	3
	Female artist, male writer	18	13	57	16	35	1
	Male artist, female writer	2	0	7	2	1	0
	Male artist, male writer	107	34	263	45	123	6

Table A4. Results of Mann-Whitney U Tests Comparing Male and Female Songwriters

Variable	Male Songwriters Median (Mean ± SD)	Female Songwriters Median (Mean ± SD)	*U*	*p*
Emotion-laden word rate (per 100 words)	9.272 (10.029 ± 4.652)	9.449 (10.110 ±4.269)	135,253.5	0.568
Primary Emotion Confidence Score	0.962 (0.869 ±0.170)	0.960 (0.874 ±0.158)	139,701.0	0.865
Sadness	0.013 (0.235 ±0.376)	0.016 (0.248 ±0.387)	134,381.0	0.474
Anger	0.010 (0.171 ±0.301)	0.008 (0.138 ±0.281)	146,126.0	0.215
Joy	0.192 (0.422 ±0.433)	0.229 (0.439 ±0.431)	134,307.0	0.466
Love	0.002 (0.094 ±0.258)	0.002 (0.089 ±0.244)	135,378.0	0.582
Fear	0.002 (0.068 ±0.210)	0.002 (0.062 ±0.196)	139,253.0	0.924
Surprise	0.001 (0.010 ±0.074)	0.001 (0.024 ±0.126)	140,924.0	0.709

Table A5. Results of Mann-Whitney U Tests Comparing Male and Female Artists

Variable	Male Artists Median (Mean ± SD)	Female Artists Median (Mean ± SD)	*U*	*p*
Emotion-laden word rate (per 100 words)	9.350 (9.997 ±4.510)	9.708 (10.414 ±4.800)	768,974.5	0.058*
Primary Emotion Confidence Score	0.955 (0.863 ±0.172)	0.974 (0.878 ±0.166)	752,973.0	0.006**
Sadness	0.015 (0.228 ±0.369)	0.012 (0.252 ±0.389)	815,979.5	0.543
Anger	0.014 (0.194 ±0.316)	0.008 (0.150 ±0.292)	879,195.0	<0.001***
Joy	0.248 (0.433 ±0.430)	0.221 (0.435 ±0.437)	807,452.5	0.878
Love	0.003 (0.081 ±0.234)	0.002 (0.094 ±0.255)	820,081.0	0.409
Fear	0.002 (0.056 ±0.189)	0.002 (0.057 ±0.194)	854,383.0	0.008**
Surprise	0.001 (0.009 ±0.070)	0.001 (0.011 ±0.081)	848,844.5	0.018**