

What Are Iron Maiden, Taylor Swift, and Pharrell Williams Singing About?

A Primer on Using Text Mining Techniques to Analyze Music Lyrics *

Walter Guillioli

April 28, 2025

Introduction

Do bands like Iron Maiden only sing about the devil? Does Taylor Swift only sing about love? Does Pharrell Williams only sing about happiness? Is there a way to apply text-mining techniques to understand song lyrics without reading them all?

I am deeply passionate about music and data science, and these questions sparked the motivation for this research. The purpose of this paper is to apply key text analytics techniques to explore and compare the lyrics of different artists. I intentionally selected Iron Maiden — my favorite band since I was 10 years old — alongside very different artists to create a contrasting analysis.

In this article, I apply basic text-mining techniques to these lyrics. I begin by identifying the most common words and then examine the emotional tone of these words. Next, I explore how individual words contribute to the overall emotional landscape of an album. Finally, I investigate how words combine together to provide deeper insights, highlighting some of the limitations of relying solely on single-word or unigram analysis.

This file is a summary. To view the code and data, see this Github repo.

Data and Methods

The dataset consists of lyrics from four different albums, downloaded from azlyrics.com. Each song was saved into a text file and subsequently loaded into a data frame for analysis. The four albums, comprising a total of 48 songs, include two albums by Iron Maiden, one by Taylor Swift, and one by Pharrell Williams. Refer to Table 1 for an overview.

The data captured for each song includes the artist's name, album name, song number, song title, and the complete lyrics. See Table 2.

The lyrics were then tokenized — first into lines, and subsequently into words. Common words that typically do not add semantic value (stop words) were removed from the dataset.

Most of the analysis was conducted at the word level. I began by counting the most frequent words in each album to gain an initial understanding of the language commonly used by each artist. This was approached in two ways: 1. A simple count of word frequencies. 2. The use of the TF-IDF (Term Frequency–Inverse Document Frequency) statistic that highlights words that are significant to a particular album but infrequent across others.

*The data and scripts posted here: https://github.com/wguillioli/music_lyrics

Each word was then mapped to a sentiment dictionary to classify the emotional tone associated with it. For example, the word happy is associated with anticipation, joy, positivity, and trust, while unhappy maps to anger, disgust, negativity, and sadness.

As will be shown in the results, all artists share a particular emotional theme — fear. We explored this further by identifying the top words contributing to this sentiment for each artist.

Additionally, words were categorized as either positive or negative. This served two purposes: first, to identify the top positive and negative words for each artist; and second, to calculate a negativity ratio at the song level. This ratio is defined as the number of negative words divided by the total number of words in the song, where a higher value indicates a more negative song. We also chart the evolution of sentiment across the album’s tracklist to understand how the listener’s emotional journey progresses.

Finally, we briefly explored lyrics at the level of word pairs (bigrams), addressing some of the limitations inherent in single-word (unigram) analysis.

Table 1: Number of Songs and Words per Album

Album	Songs	Words
Iron Maiden: Number of the Beast	9	800
Iron Maiden: X Factor	11	871
Pharrell Williams: GIRL	12	1437
Taylor Swift: 1989	16	2000

Table 2: List of Songs per Album

Taylor Swift: 1989	Pharrell Williams: GIRL	Iron Maiden: Sign of the Cross	Iron Maiden: Number of the Beast
Welcome to NY	Marilyn Monroe	Sign of the Cross	Invaders
Blank Space	Brand New	Lord of the Flies	Children Of the Damned
Style	Hunter	Man on the Ledge	The Prisoner
Out in the Woods	Gush	Fortunes of War	22 Acacia Avenue
All you had to...	Happy	Look for the Truth	Number Of The Beast
Shake it off	Come Get It Bae	The Aftermath	Run to the Hills
I wish you would	Gust of Wind	Judgment of Heaven	Gangland
Bad blood	Lost Queen	Blood on the...	Hallowed be thy Name
Wildest Dreams	Freq	Edge of Darkness	Total Eclipse
How you get the girl	Know who you are	2AM	
This love	It Girl	The Unbeliever	
I know places	Smile		
Clean			
Wonderland			
You are in love			
New romantics			

Results

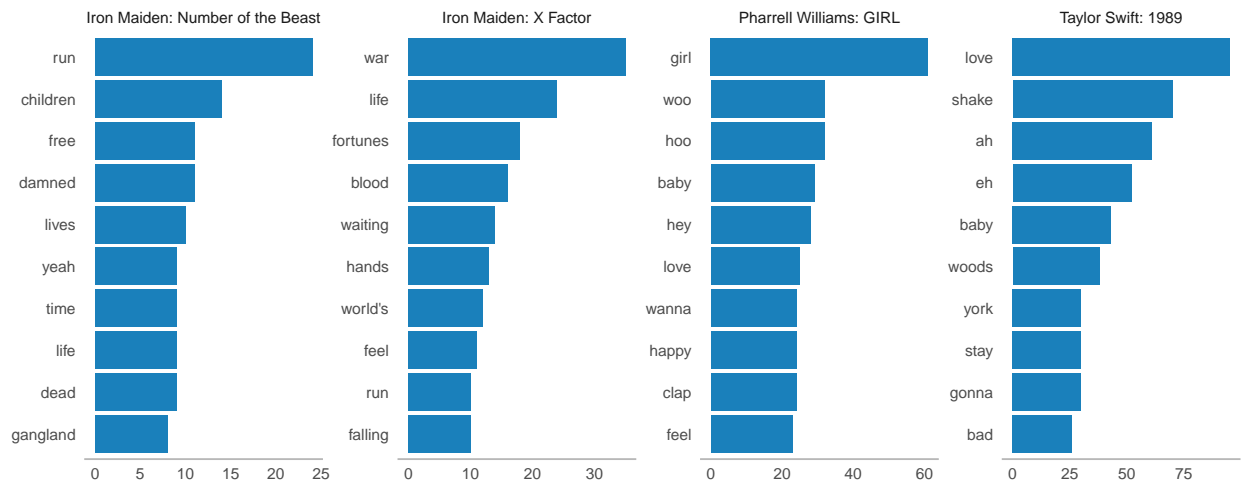
Top Words Used by Artist

As Figure 1 illustrates, Iron Maiden tends to use stronger and more negative language in their lyrics. In one album, prominent words include war, blood, and failing, while in the other, words such as beast, damned, and dead are more common.

Pharrell Williams, by contrast, predominantly expresses themes of happiness, often singing about girls and using energetic expressions such as woo-hoo to convey excitement.

Finally, Taylor Swift shows a strong inclination toward singing about love, while also incorporating more informal language, with frequent use of words like baby, ah, and eh.

Figure 1: Count of Most Used Words per Album



Sentiment Evolution in the Album

How are emotions experienced throughout the albums? Figure 2 illustrates the sentiment evolution across some of the albums as the tracks progress line by line.

Several observations are worth highlighting. First, the number of lines varies significantly among the artists: Iron Maiden's songs tend to have fewer lines compared to those of Pharrell Williams and Taylor Swift, indicating a more concise lyrical style.

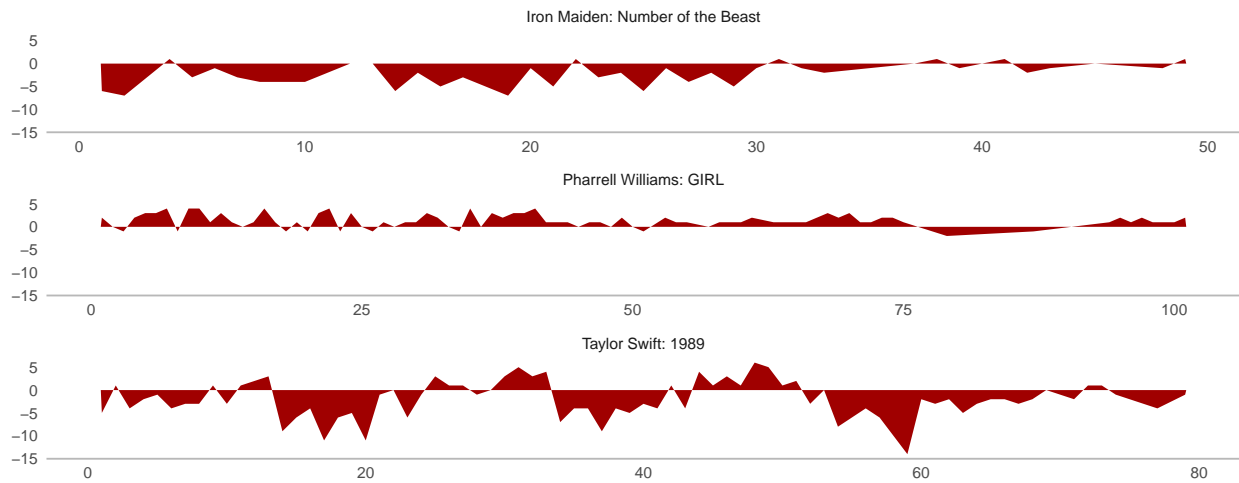
More importantly, the sentiment patterns reveal distinct differences across the artists. Iron Maiden maintains a predominantly negative sentiment throughout their albums. In contrast, Pharrell Williams consistently exhibits a positive emotional tone. Taylor Swift, however, alternates between positive and negative sentiments, reflecting a more dynamic emotional journey across her album.

Main Feelings From each Album

While analyzing the top words and sentiment evolution provides an initial understanding of the lyrical themes, it is only the beginning. The next step is to determine whether the artists predominantly sing about happiness, sadness, anger, or other emotions. As shown in Figure 3 for a sample of the albums, several key patterns emerge.

First, fear appears as a common emotion across all artists and albums, a point that will be explored further in the next section.

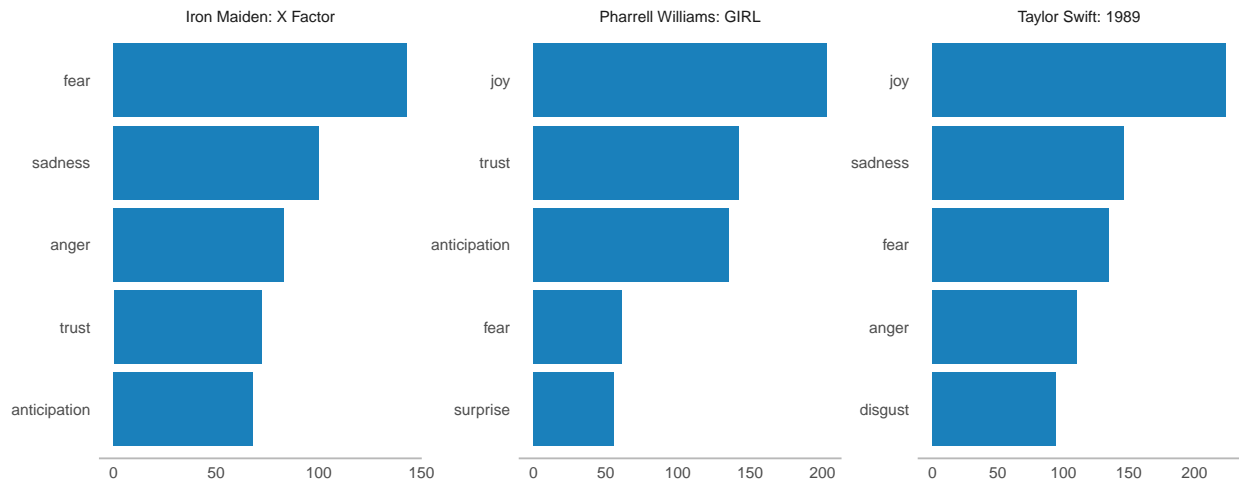
Figure 2: Sentiment Evolution From Start to Finish per Album



For Iron Maiden, the predominant emotions are negative, particularly anger and sadness. In contrast, Pharrell Williams’ lyrics are strongly associated with positive emotions such as happiness, joy, and trust. Taylor Swift presents a blend, often mixing joyful themes with sadness.

These findings are consistent with the sentiment evolution trends observed in the previous section.

Figure 3: Top 5 Emotions Based on Lyrics per Album



What Are They Afraid of?

As observed in the previous section, fear emerges as a common sentiment across all albums. A natural question that follows is: fear of what? Interestingly, it varies significantly between artists.

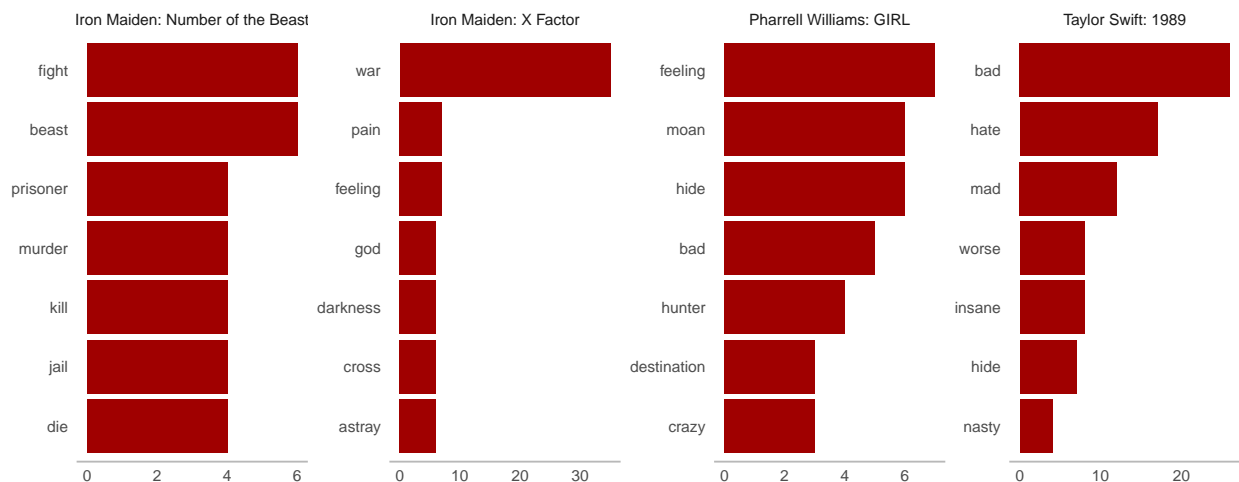
For Iron Maiden, fear is predominantly associated with violent or dark imagery, with top words including fight, beast, murder, kill, jail, and war.

Taylor Swift’s lyrics, on the other hand, center around relational fears, using words such as bad, mad, and hate.

Pharrell Williams’ results are somewhat more ambiguous and even humorous. Notably, moan appears among the top fear-associated words. While moan can carry a fearful connotation, within the context of his lyrics it is more likely related to sexual themes rather than fear. Similarly, the word hunter is flagged as negative by the sentiment dictionary; however, in his lyrics, it refers to “hunting for love,” which is not inherently negative.

These examples highlight some of the limitations of unigram-based text mining, where single words are analyzed without context. A more detailed discussion of these limitations will be presented in a later section.

Figure 4: Count of the Most Used Words in the Fear Category per Album



Positive and Negative Words

To further explore the lyrics, we examined the most positive and negative words for each album. Two of these albums are displayed in Figure 5.

In the case of Iron Maiden, free emerges as one of the top positive words. While it is indeed categorized as positive, a closer reading of the lyrics reveals that it often refers to “running free” from the police, which is not necessarily a purely positive theme. Similarly, hallowed appears as a top positive word, but it is primarily used in the song Hallowed Be Thy Name, which tells the story of a prisoner awaiting execution. Thus, although the word carries a positive connotation, the surrounding context is far more somber. In contrast, the negative words identified align well with expectations, including terms such as damned, dead, and burning.

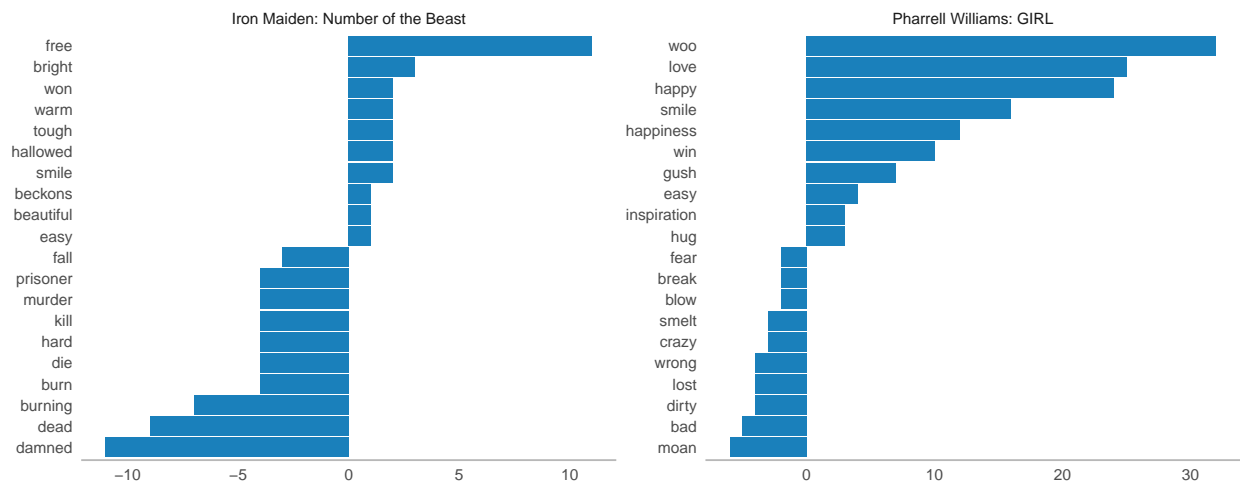
For Pharrell Williams, the results highlight a strong emphasis on positivity, with woo (as part of woo-hoo) standing out among the top positive words. As expected, the number of negative words is significantly lower in comparison.

In the case of Taylor Swift (not shown), the word love naturally leads the list of positive words. Regarding negative words, most align with expected emotional tones; however, one example worth noting is shake. Although shake is classified as negative by the sentiment dictionary, in the context of Swift’s lyrics, it refers to “shaking off” critics and negativity while celebrating a carefree, joyful life. Thus, the word itself is not used negatively within the song.

How Negative is Each Song?

A complementary approach is to analyze the albums at the song level to better understand how each track contributes to the overall sentiment of the album. As shown in Figure 6, Iron Maiden’s songs tend to have a more negative sentiment compared to those of Pharrell Williams.

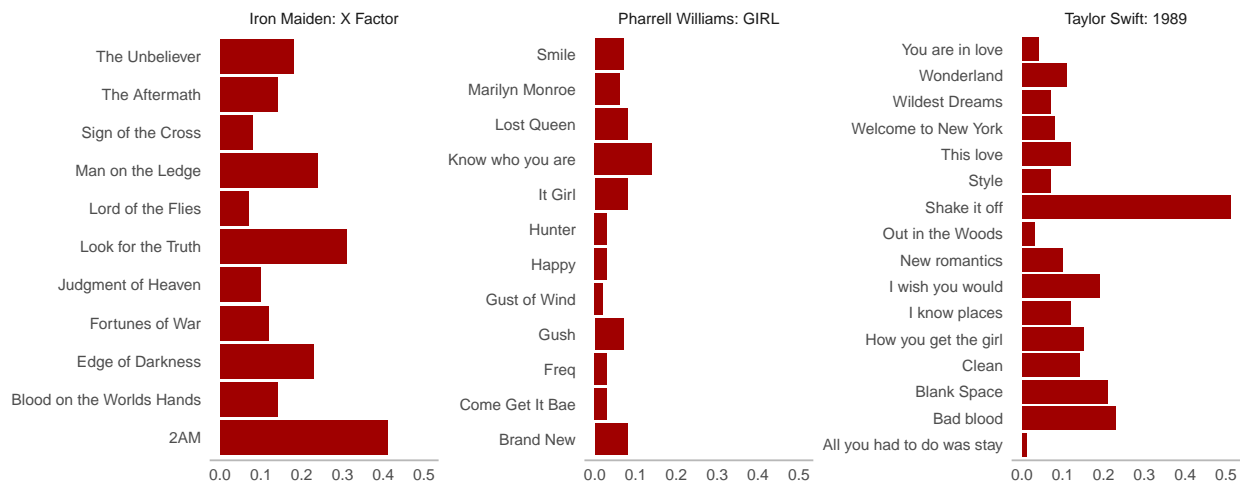
Figure 5: Count of the top 10 Most Positive and Negative Words per Album



For example, Children of the Damned addresses the plight of homeless children and their suffering, which contributes to the negative tone of the album.

In the case of Taylor Swift, one notable instance is the song Shake It Off, which is incorrectly categorized as negative due to the misinterpretation of the word shake, as discussed earlier. Another example is Bad Blood, a song about a failed relationship that transitions from love to hate. This shift in emotion aligns with the song’s negative sentiment, making it a fitting example of how sentiment analysis can reflect the thematic content of the song.

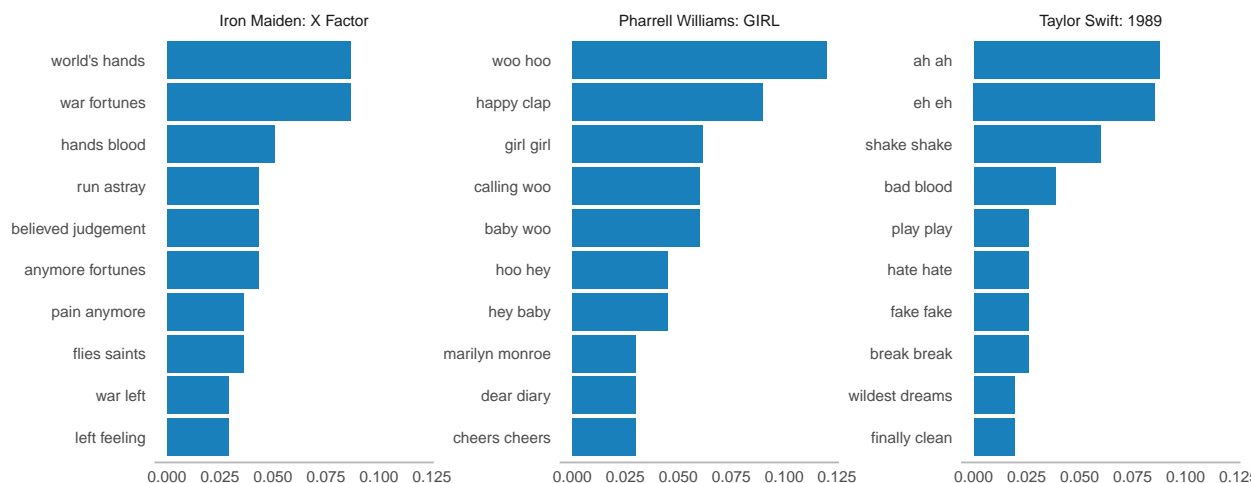
Figure 6: Negativity Score per Song



Looking Beyond One Word

Previously, we discussed text analysis primarily at the individual word level. While this approach is useful, it has certain limitations, as previously highlighted. An obvious next step is to examine the lyrics in terms of word pairs or bigrams, which are shown in Figure 7.

Figure 7: Top 10 Pairs of Words Used in Each Album



A few key points are worth highlighting here. For Iron Maiden, the word fortune is classified as positive, but in context, it is often used in association with war, referring to the negative consequences of conflict (e.g., “the fortunes of war”).

For Pharrell Williams, positivity is emphasized by word pairings such as happy and clap, as well as cheers and cheers, which reinforce the celebratory tone of his lyrics.

Taylor Swift, on the other hand, tends to use simpler and more repetitive phrasing in her lyrics. Notable examples include the frequent repetition of words like ah ah, eh eh, and fake fake, which contribute to the conversational and emotionally charged nature of her songs.

Mapping Words as Networks

Furthermore, we explore how words interconnect in the lyrics. While the size of song lyrics is smaller compared to a book, the results are still insightful and offer interesting perspectives.

Due to space constraints, we only present the network for two albums. We observe that the more elaborate combinations of words provide valuable insights. For instance, for Iron Maiden the word run is connected to several terms such as hills and lives. This makes sense within the context of their song Run to the Hills, which describes the act of running to the hills to escape from ruthless conquerors.

When contrasting these results with Figure 9, we observe that Taylor Swift’s word networks revolve around more personal and adventurous themes, such as “wanna hug”, “lost queen”, and other similarly emotive expressions.

Impact of Negated Words

Finally, it is crucial to consider the impact of negated words on sentiment analysis. For example, the word happy is classified as positive when analyzed in isolation, but its meaning can change significantly in

certain contexts. For instance, “I am happy” is positive, while “I am not happy” should be interpreted as negative. However, sentiment analysis tools often misclassify such negated phrases by treating happy as positive regardless of context. Therefore, it’s important to identify words that are preceded by negators such as “not”, “no”, “never”, or “without” and examine how they may mislead sentiment scores.

Figure 8: Network of Top Bi-grams for Iron Maiden’s Number of the Beast Album

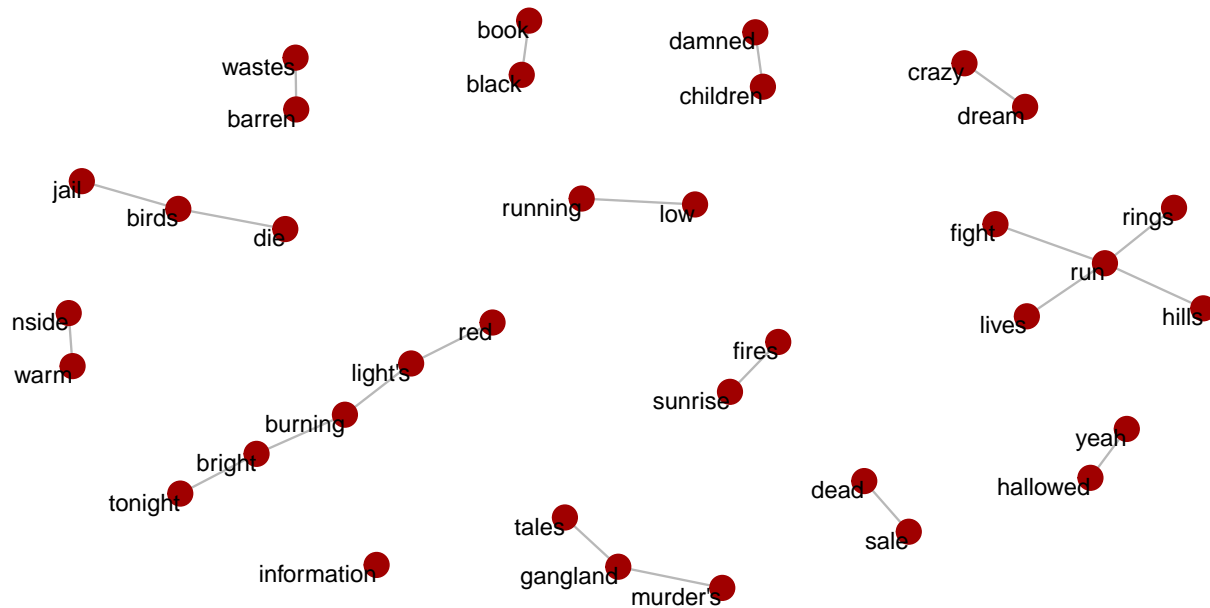
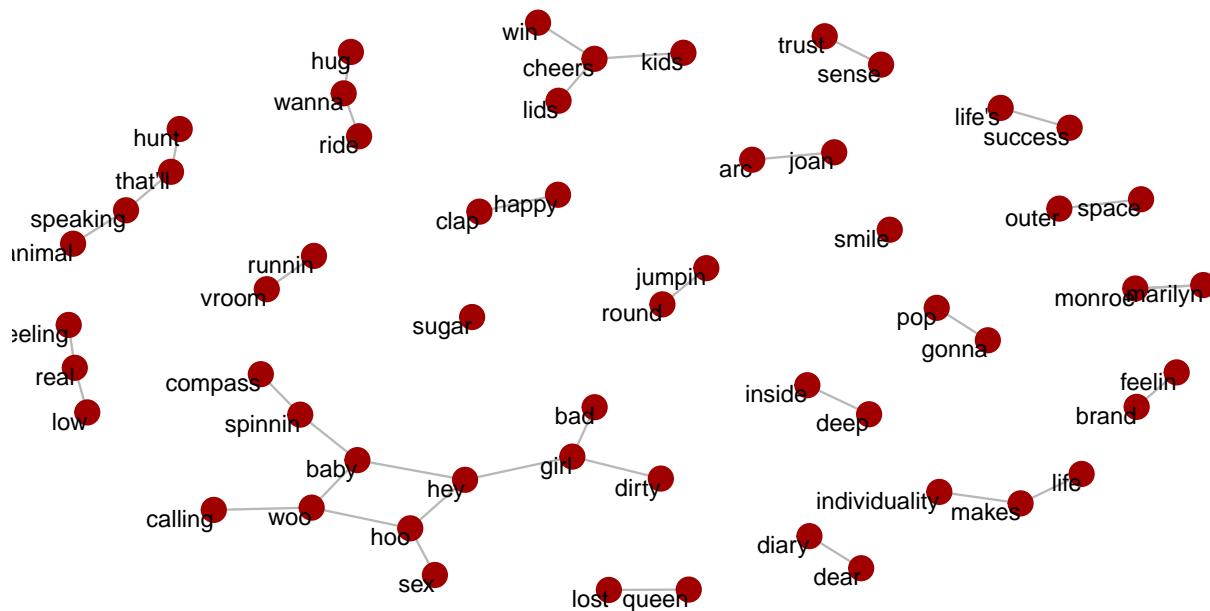
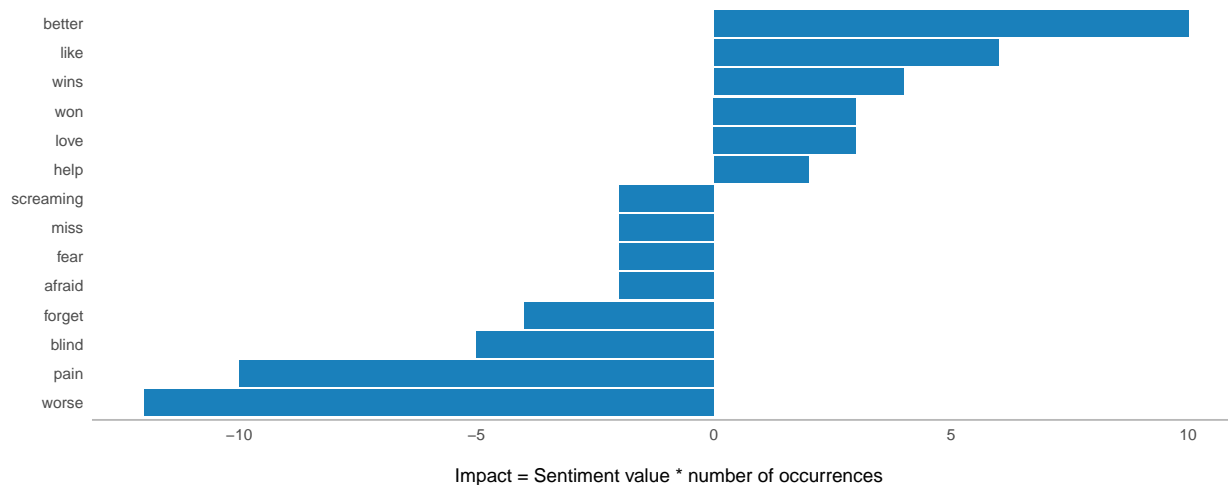


Figure 9: Network of Top Bi-grams for Taylor Swift’s 1989 Album



As shown in Figure 10, the word *worse* is classified as highly negative, but in the context of “I am not worse”, it would actually convey a more positive sentiment. This issue applies to other words as well. A clear example is the word *better*, which is positive on its own, but if preceded by “not”, it takes on a negative meaning.

Figure 10: Top Words Misclassified by Preceded Negated Words in Analysis



Conclusions

1. *Text mining techniques* prove to be highly effective for processing song lyrics. Basic methods such as word counting, tf-idf, unigram analysis, and sentiment analysis can provide valuable insights into the primary vocabulary and emotional tone of each artist.
2. *However, there are notable limitations* to consider. The presence of negated words can significantly distort sentiment analysis results and should not be overlooked. Therefore, an analysis using bigrams is recommended to capture more nuanced relationships between words and their context.
3. *Lyrics differ by genre*. Iron Maiden's lyrics tend to be more intricate, concise, and meaning-driven, while popular music (such as pop) often relies on simpler, more repetitive language that may prioritize catchiness over depth.
4. *Iron Maiden's lyrics are predominantly negative*, not in a purely pessimistic sense, but rather as a reflection of profound societal issues. Their songs often explore themes such as war, suffering, and existential questions, presenting a more complex emotional landscape.

References

1. AZLyrics. (n.d.). AZLyrics – Song lyrics from A to Z. Retrieved April 7, 2025, from <https://www.azlyrics.com/>
2. Bird, S., Klein, E., & Loper, E. (n.d.). Natural Language Toolkit. Nltk.org. Retrieved April 7, 2025, from <https://www.nltk.org/>
3. Datanovia. (2021, October 1). ggplot2 themes gallery. Datanovia. <https://www.datanovia.com/en/blog/ggplot-themes-gallery/>
4. Holtz, Y. (n.d.). Reorder a variable in ggplot2. The R Graph Gallery. Retrieved April 7, 2025, from <https://r-graph-gallery.com/267-reorder-a-variable-in-ggplot2.html>
5. Rafael, F. R. (2021, March 7). Sentiment analysis with tidytext. RPub. <https://rpubs.com/rafrys/723764>

6. Silge, J., & Robinson, D. (n.d.). Text Mining with R. Tidytextmining.com. Retrieved April 7, 2025, from <https://www.tidytextmining.com/>
7. Wickham, H. (n.d.). ggplot2: Elegant graphics for data analysis. Ggplot2-book.org. Retrieved April 7, 2025, from <https://ggplot2-book.org/>