1. Results

1.1 Analyze Raw Data

By comparing the percentage of death-songs and sampled-all-songs, the following observations can be made:

- Genre Impact: Death-related words have more possibilities to be mentioned in the music genres such as rock, reggae and hip hop compared to country, pop, jazz and blues.
- Temporal Trends: Over time, the number of songs expressing death has gradually increased, and notably, there has been a substantial surge in recent years.
- Theme Focus: Songs addressing death mostly focus on topics related to violence, with a smaller proportion focusing on obscene.

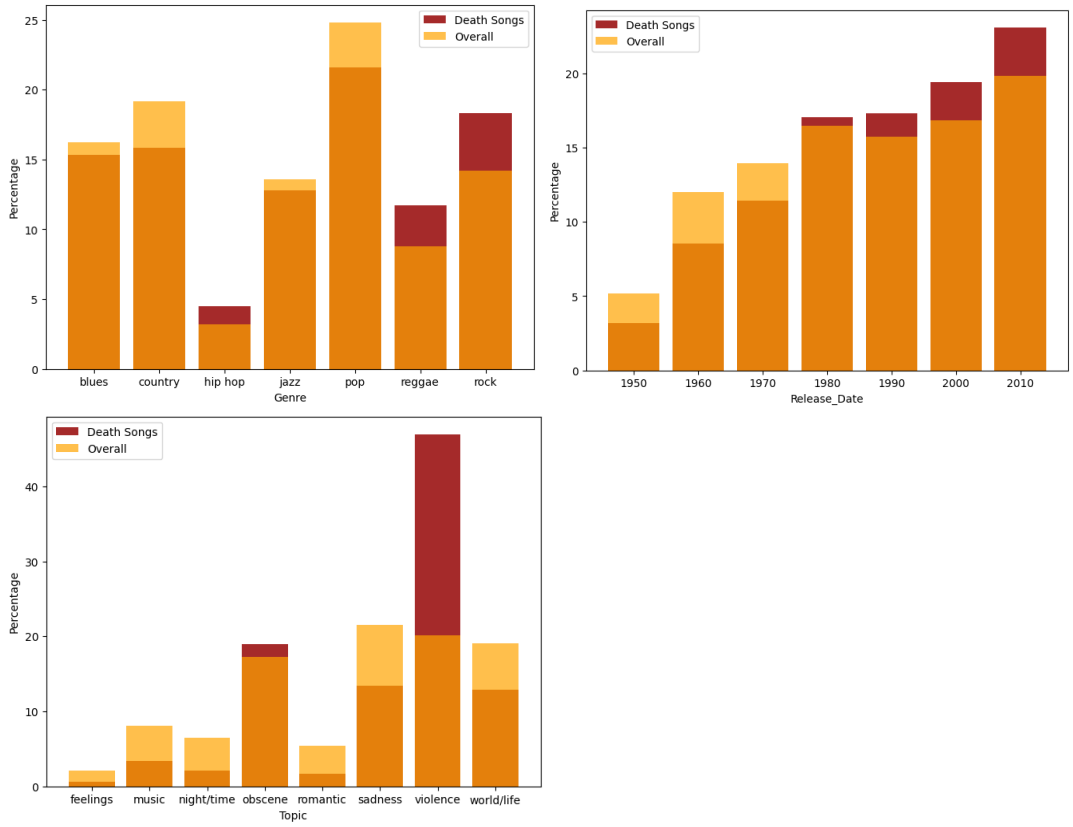

Figure1&2&3: Percentage of Death-Songs and Whole-Sampled-Songs by Feature 'Genre', 'Released_date', and 'Topic'

1.2 Topic Modelling - LSA

The results of LSA in both datasets reveal distinct patterns:

- Death-Songs LSA: The analysis of death-songs showcases more aggressive and negative words, such as 'life', 'fight', 'money', 'heart', and 'smoke', which are closely associated with themes of death. Remarkably, we observed that death was also linked to racially offensive words.
- All-Songs LSA: Within the sampled dataset, while negative words are still present, there is a notable presence of positive words like 'love', 'home', 'good', and 'dream.' , which are not as prominent in death-songs.

Figure 4: The LSA result of Death-Songs

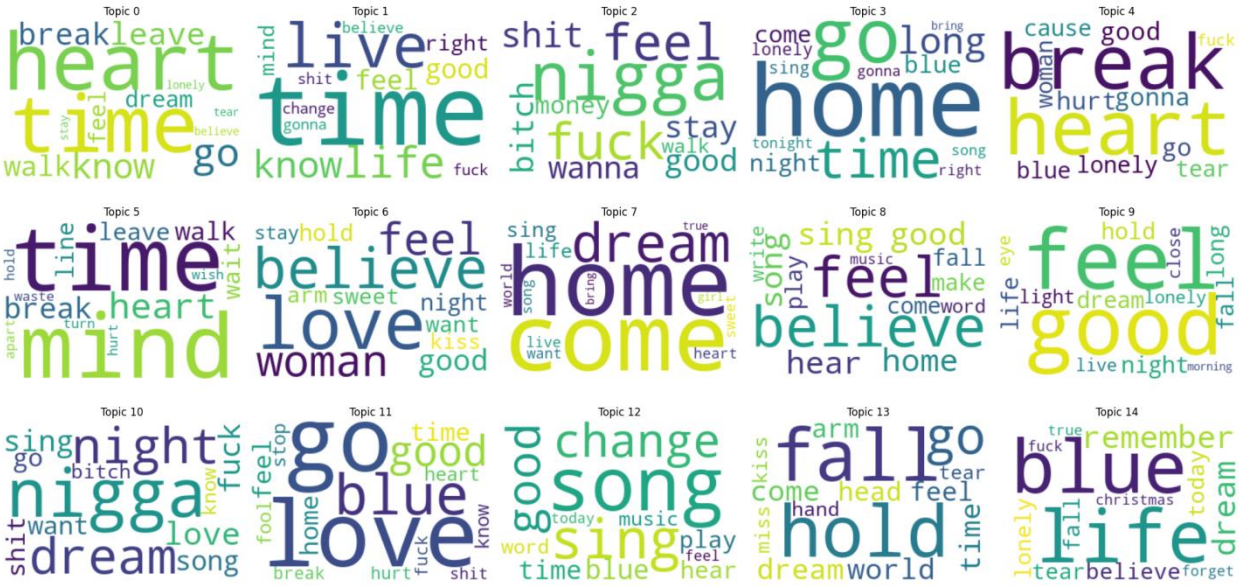

Figure 5: The LSA result of All Songs

1.3 Topic Modelling - LDA

The results of LDA in both datasets reveal distinct patterns:

- Death-Songs LDA: Similar to LSA, LDA exhibits topics featuring words like ‘life’, ‘money’, ‘time’, and ‘fight’. Additionally, certain topics, such as

Topic 12, reveals a diverse array of words such as ‘crawl’, ‘worm’, ‘fear’, ‘fever’, creating an evocative association with the concept of death. We also find that aggressive words are grouped together in Topic 4.

- All-Songs LDA: In the sampled dataset, besides aggressive words grouped under Topic14, various lyric themes emerge. Some topics like ‘fight’ and ‘life’ can be connected with death-lyrics. Others like ‘love’ and ‘true’ seem more prevalent in non-death-lyrics.

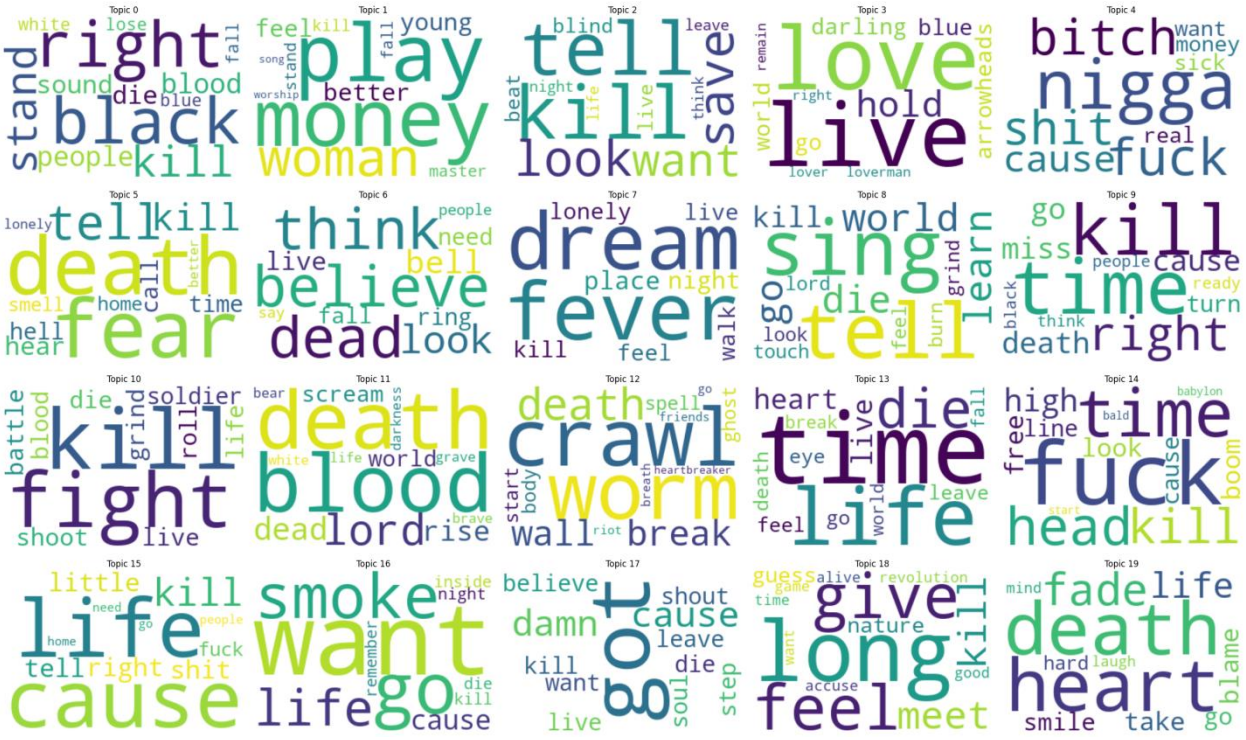

Figure 6: The LDA result of Death-Songs

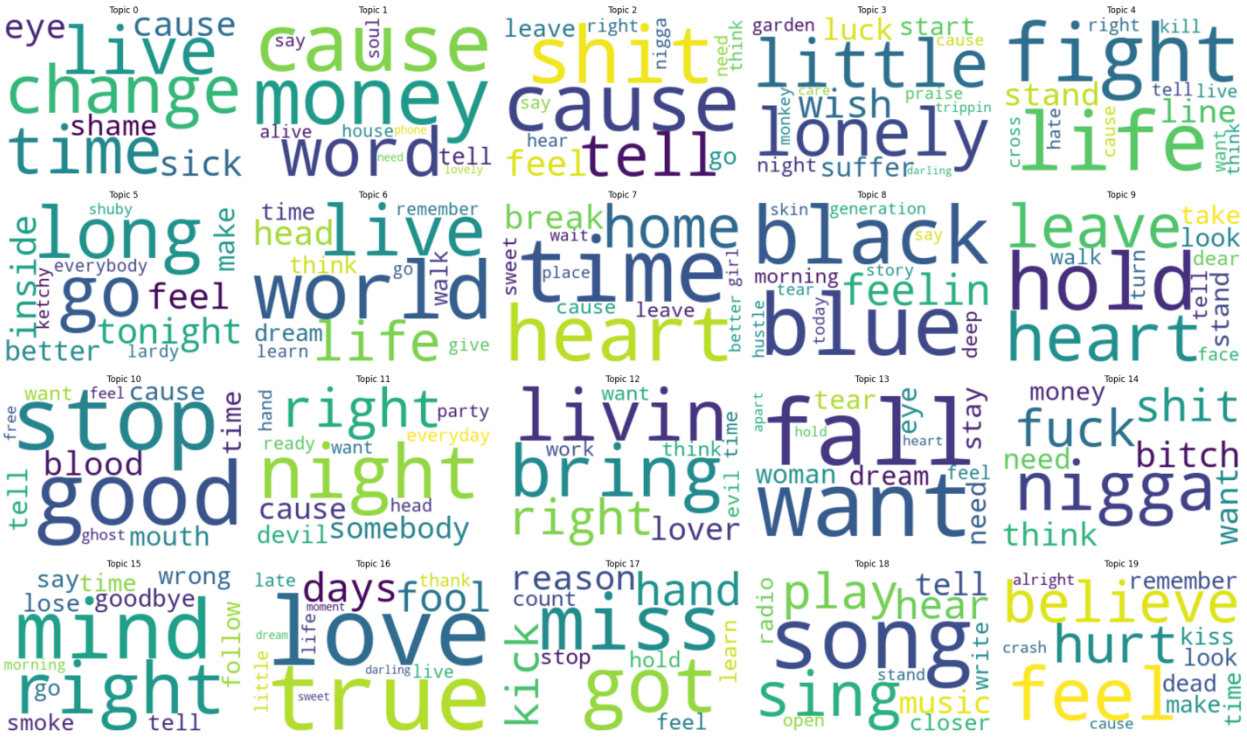

Figure 7: The LDA result of Sample-Songs

1.4 Use Classified lyrics for Topic Modelling

After classifying the death-lyrics by genre and stopping the same top words like 'kill', 'death', 'die', 'come', 'know', 'time', 'life', 'live', we can find that:

- Aggressive Themes: Hip hop, rock, jazz, and reggae genres exhibit more aggressive language in their death-related lyrics.
- Genre-Specific Topics: Jazz often explores themes related to the 'world', and Country genres frequently touch upon topics associated with the 'heart.'
- Rock Genre Insights: Based on the findings before, rock stands out for its connection to death-related topics. Noteworthy topics in rock lyrics include 'fight', 'world', 'heart', among others.

Figure 8: The LDA result of Death-Songs Classified by Genre

After applying the same methodology mentioned earlier, I delved into the topics of death-related lyrics across various release dates. The findings are as follows:

- Aggressive Themes Over Time: The expression of death in lyrics has become more aggressive as time progresses. This trend is particularly evident under the topic of the 2010s, aligning with the decade's peak in songs expressing death.
- Evolution of Themes: During the 1950s to 1970s, the word 'heart' frequently accompanied discussions about death in lyrics. As time advanced, the thematic focus expanded to include concepts such as ‘live’, ‘feel’, and ‘world’.

Figure 9: The LDA result of Death-Songs Classified by Release_date

Also applying the same methodology into the death-related-lyrics classified by topic.

The findings are as follows:

- Violent and Obscene Topics: Confirming the observations from Part 1, where most death-related songs were identified in topic violence and obscenity, the analysis here reveals a notable prevalence of aggressive and death-associated words within these topics.
- Limited Impact in Romantic Topics: In contrast, the topic related to romance shows minimal influence in this analysis. Death-related words and aggressive themes appear to be less prominent in the context of romantic lyrics.

Figure 10: The LDA result of Death-Songs Classified by Topic

2. Discussion

Based on the purpose of project and the techniques used in project, I'll discuss three following sections:

2.1 Evaluation and Comparison between LSA and LDA

During the exploration of LSA, adjusting the number of topics revealed a limitation that higher numbers led to repeated content, reducing interpretability. Particularly, the concentration on negative words related to death were more evident through LSA, instead of positive words like 'love', 'good' and 'home' created in all-lyrics dataset.

Contrasting with LSA, LDA demonstrated a strength in producing diverse topics. Besides the same words like 'life' and 'fight', it generated unexpected terms like

‘crawl’, ‘worm’ and ‘fever’, providing a nuanced understanding of death-related themes. And with the increase of number of topics, more different content appears. However, the multitude of topics in LDA made it more challenging to tell the difference between death-lyrics and all-lyrics dataset.

Drawing insights from both my experimentation and relevant literature[3], I've distilled the comparisons between LSA and LDA into the following conclusions:

- LSA is a topic modelling technique which always capture the relationship between the documents and the terms they contain. Illustrated by our results, LSA revealed words like ‘life’, ‘heart’, ‘fight’, ‘night’, signifying a semantic resonance with death.
- LDA, on the other hand, excels in identifying topics that documents belong to based on the word contained within them. Instead of discovering relationships, LDA seeks to discover the underlying topics in a corpus of text. In the death-lyrics experiment, LDA unfolded a diverse array of topics like ‘crawl, worm’, ‘fever, lonely’, ‘battle, soldier’, which generate a narrative and evoke sentiments akin to the theme of death.
- Under the purpose of exploring the topics expressing death, LDA emerged as the more effective technique, producing richer and more generative topics related to death.
- When the aim was to explore disparities between death-lyrics and non-death-lyrics, LSA exhibited a more pronounced effectiveness.

2.2 Bias on death in song lyrics

From the result we get, it underscore a consistent association between death and aggressive, negative words. This inevitably raises a dialectical question that Why do people tend to think that death is a bad thing?

Some perspectives posit that the negativity surrounding death comes from its inherent nature. “It’s bad because it deprives us of good things we would otherwise have had.” Relatively, certain philosophical viewpoints assert that where life can’t exist without death, which allows individuals to use death as a reminder to enjoy life.

Undoubtedly, this will always be a philosophically complex question. The encouraging aspect, as revealed through song lyrics, is that more individuals are engaging in conversations about death, which indicates a growing awareness and willingness to explore the multifaceted nature of death.

2.3 Future plan

Death is a very tricky theme and what I explored in this project is just a little start. The following are proposed steps to enhance the project:

- Web scraping the country of each song: Implement web scraping techniques to gather information on the country for each song, which can contribute valuable insights into how cultural backgrounds may influence the expressions on death in song lyrics.

- Use RNNs to Generate death-lyrics: Explore the implementation of Recurrent Neural Networks (RNNs) to generate a series of song lyrics based on different genres, release dates and topics.
- Try to build a Chatbot like ELIZA to talk about death: Tailor a chatbot to engage users in discussions about death, posing questions and facilitating dialectics to stimulate users to reflect more deeply on the philosophical aspects of death.