



# **Song Topic Cluster Analysis: Unveiling Musical Trends**

Beining Zhang



# Agenda

- Data Extraction
- Data Cleaning
- Exploratory Data Analysis (EDA)
- Topic Clustering
- Interpretation of Clusters
- Result
- Limitation
- Future Work



# Data Extraction

Folder with txt files -> A DataFrame with columns of songs names and lyrics

	Songs	Lyrics
0	Anna-go-to-him	Anna, you come and ask me girl\nTo set you fre...
1	yesterday	Yesterday, all my troubles seemed so far away\...
2	i-need-you	You don't realize how much I need you\nLove yo...
3	the-ballad-of-john-and-yoko	Standing in the dock at Southampton\ntrying to...
4	one-after-909	My baby says she's traveling\nnon the one after...



# Data Cleaning

Removed blank files and songs with duplicated lyrics.

Number of songs before cleaning: 279

Number of songs after cleaning: 217

Number of blank files: 58

Number of duplicated songs: 4

Total number of songs removed: 62



# Exploratory Data Analysis (EDA)

How many of the songs feature at least one pair of lines that rhyme?

Define rhyme: the last words in the consecutive two lines of lyrics are ending with the same 2 letters.

Method: For each song's lyrics, split into lines and extract the last words from every two consecutive lines and check whether the two letters in the last words match.

Data preprocessing: Lowercase the lyrics.

Number of songs with at least one pair of rhyming lines: 199



# Exploratory Data Analysis (EDA)

How many of the songs feature the song name (found in the file name) in the song lyrics?

Method: For each song, search in its lyrics whether the song name is in it.

Data Preprocessing: The dash in song names are replaced with space. The lyrics are lowercased, removed whitespace, and punctuations.

Number of songs that feature the song name in the lyrics: 186

Cleaned_Songs	Cleaned_lyrics	Song Name in Lyrics
yesterday	yesterday all my troubles seemed so far away n...	True
i need you	you dont realize how much i need you lve you ...	True
one after 909	my baby says shes traveling on the one after 9..	True
shes a woman	my love dont give me presents i know that shes...	True
three cool cats	three cool cats three cool cats are coming up ...	True

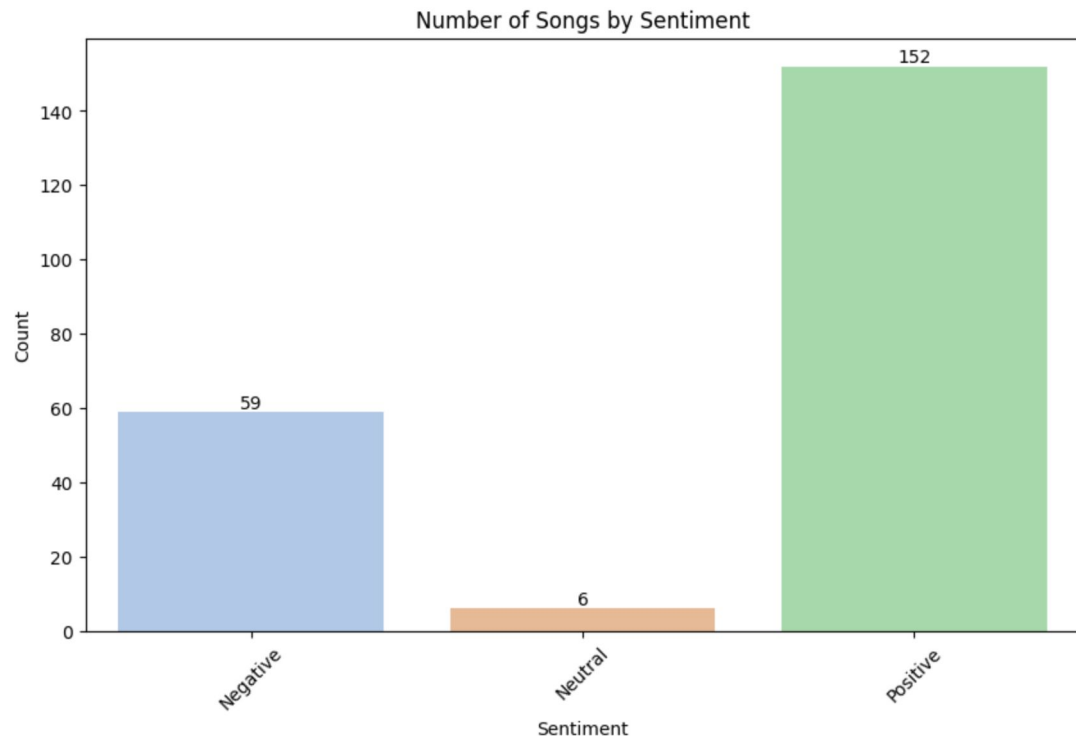


# Exploratory Data Analysis (EDA)

## Sentiment Analysis

Method: Used `SentimentIntensityAnalyzer` to score the lyrics and classify them as “Negative”, “Neutral”, or “Positive” (using  $\pm 0.05$  thresholds).

Data Preprocessing: Further cleaned lyrics by tokenizing and removing stopwords.





[illegible][illegible]



# Topic Clustering

Method: Latent Dirichlet Allocation (LDA) topic modeling

Data Preprocessing:

- Lemmatization
- Create a dictionary representation of the documents
- Convert the tokenized documents into a Document-Term Matrix (DTM)

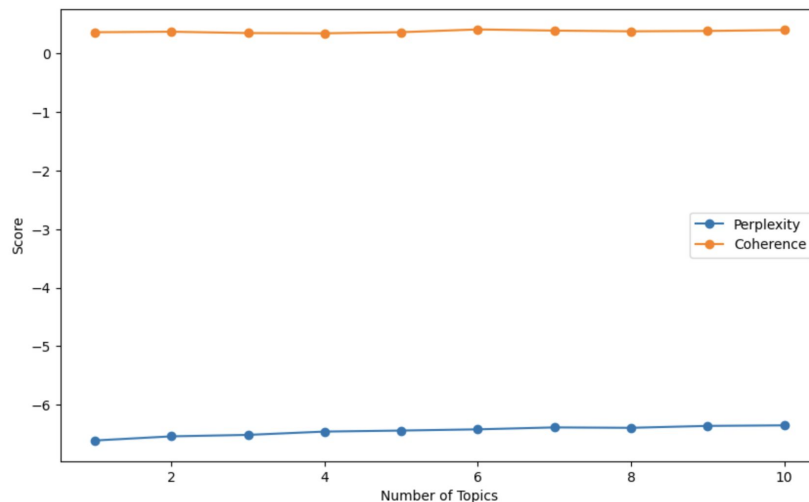
Number of topics: 5 (according to the Perplexity and Coherence scores)

# Choose number of topics

Lower perplexity values indicate a better fit to the data. Higher coherence scores indicate that the topics are more interpretable and coherent. In our case, 5 or 6 seems suitable.

- 5 topics: It has a relatively low perplexity score and a good coherence score.
- 6 topics: It has a slightly higher perplexity score but the highest coherence score, indicating more interpretable topics.

```
Nmber of topics: 1, Perplexity score: -6.607932699363433, Coherence score: 0.35940413508114166
Nmber of topics: 2, Perplexity score: -6.53728265365617, Coherence score: 0.36852994107013926
Nmber of topics: 3, Perplexity score: -6.512135686049541, Coherence score: 0.34551849415797603
Nmber of topics: 4, Perplexity score: -6.455839594386056, Coherence score: 0.34144181985679906
Nmber of topics: 5, Perplexity score: -6.438364191590927, Coherence score: 0.3683690789544701
Nmber of topics: 6, Perplexity score: -6.417471866162751, Coherence score: 0.40753132169080336
Nmber of topics: 7, Perplexity score: -6.384586508711286, Coherence score: 0.38761994315402554
Nmber of topics: 8, Perplexity score: -6.391590230861351, Coherence score: 0.3750698240007032
Nmber of topics: 9, Perplexity score: -6.358783049099269, Coherence score: 0.3817416989714838
Nmber of topics: 10, Perplexity score: -6.349536488604738, Coherence score: 0.39694951816645435
```





# Interpretation of Clusters

## **Topic 1: Love and Affection**

This topic seems to revolve around themes of love and affection. Words like "love," "im," "know," "want," "got," and "girl" suggest a strong emphasis on love and romantic relationships. It conveys feelings of desire and affection.

## **Topic 2: Actions and Expressions**

This topic appears to be about actions and expressions. Words like "na," "let," "gon," "see," and "together" suggest a focus on taking action and expressing oneself. It might involve planning or making decisions.

## **Topic 3: Assertiveness and Statements**

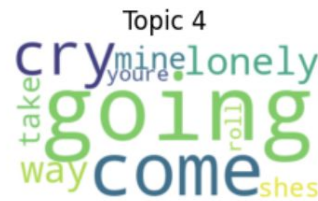
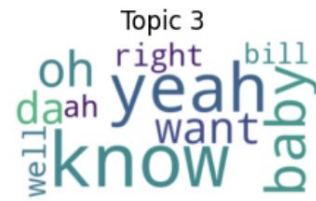
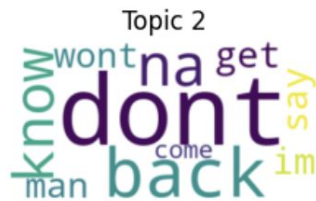
This topic seems to involve assertiveness and statements. Words like "dont," "back," "na," "get," "say," and "wont" indicate a tone of making statements or expressing one's perspective with confidence.

## **Topic 4: Positive Emotions and Affirmations**

This topic is related to positive emotions and affirmations. Words like "know," "yeah," "baby," "oh," and "want" indicate a positive and enthusiastic tone. It might involve expressions of joy, affirmation, or optimism.

## **Topic 5: Emotions and Expression**

This topic could be about emotions and expressions. Words like "going," "come," "cry," "lonely," and "mine" suggest a focus on emotions and expressions, possibly related to personal experiences and feelings.



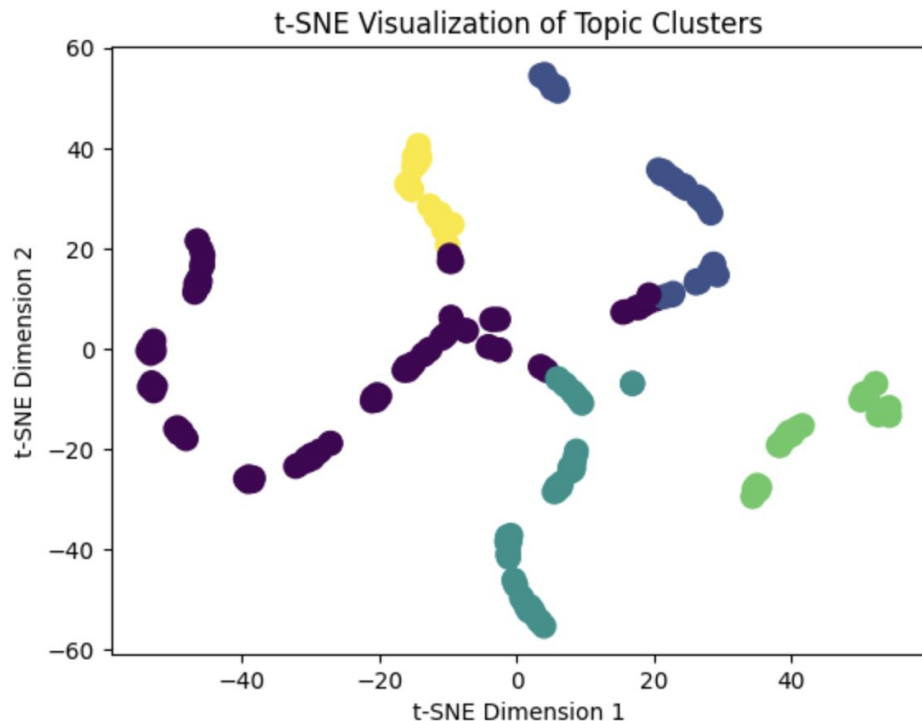
Topic 1: Love and Affection

Topic 2: Actions and Expressions

Topic 3: Assertiveness and Statements

Topic 4: Positive Emotions and Affirmations

Topic 5: Emotions and Expression





## Result

Songs	Predicted Topic
the-sunshine-arrived	0
peace-love-and-peace	0
a-very-beatles-christmas	3
love-is-all-that-s-needed	0
a-less-yellow-submarine	1

Topic 0: Love and Affection

Topic 1: Actions and Expressions

Topic 2: Assertiveness and Statements

Topic 3: Positive Emotions and Affirmations

Topic 4: Emotions and Expression



# Result

## Topic 0: Love and Affection

### The sunshine arrived:

In the morning light I feel so free  
With the sun shining bright, **just you and me**  
**Walking hand in hand**, down a road of dreams  
With every step we take, it seems like we're living in a scene  
Oh, we're walking on sunshine, and the sky is so blue  
With the breeze blowing softly, it feels like we're flying too  
And the world is so bright, **when we're together**, it's true  
Oh, we're walking on sunshine, living life like it's new  
In the fields of gold, we laugh and we play  
With the flowers and the trees, it feels like a perfect day  
We dance and we sing, to the rhythm of our hearts  
With every beat and every note, **it feels like we're never apart**  
Oh, we're walking on sunshine, and the sky is so blue  
With the breeze blowing softly, it feels like we're flying too  
And the world is so bright, when we're together, it's true  
Oh, we're walking on sunshine, living life like it's new  
**With every smile and every kiss**, it feels like a dream come true  
And every moment we spend, together it's all brand new  
Oh, we're walking on sunshine, and the sky is so blue  
With the breeze blowing softly, it feels like we're flying too  
And the world is so bright, when we're together, it's true  
Oh, we're walking on sunshine, living life like it's new  
And when the night falls and the stars shine bright  
We'll be walking on sunshine, into the endless night.



# Result

## Topic 0: Love and Affection

### Peace love and peace:

In a world that's filled with chaos and strife  
We search for a beacon, to lead us to life  
With love in our hearts, we'll conquer all fear  
With peace in our minds, we'll spread it far and near  
Peace and love, it's what we all need  
With these words, we'll plant the seeds  
Of a world that's filled with harmony and glee  
Where everyone can live in peace and love, you'll see  
From the mountains high, to the valleys below  
We'll sing out this message, for all to know  
That peace and love, is what we need the most  
And together we'll build a world that's close  
Peace and love, it's what we all need  
With these words, we'll plant the seeds  
Of a world that's filled with harmony and glee  
Where everyone can live in peace and love, you'll see  
Let us spread this message, to every land and sea  
With peace and love, we'll set our spirits free  
Peace and love, it's what we all need  
With these words, we'll plant the seeds  
Of a world that's filled with harmony and glee  
Where everyone can live in peace and love, you'll see  
With peace and love, we'll make a brighter day  
In a world that's filled with peace and love, we'll find our way.





# Result

## Topic 3: Positive Emotions and Affirmations

### A very beatles christmas:

It's that time of year, when the world comes alive  
With the sounds of joy, and the twinkling of lights  
The air is so crisp, and the snow falls down  
And the spirit of Christmas, can be felt all around  
It's the most wonderful time of the year  
With the laughter and cheer, ringing in our ears  
It's a season of love, and a season of peace  
And a time to be grateful, for all that we share and believe  
In the homes and the streets, the carolers sing  
With the songs of the season, their voices take wing  
And the children are giddy, with excitement and glee  
As they dream of the gifts, under the tree  
It's the most wonderful time of the year  
With the laughter and cheer, ringing in our ears  
It's a season of love, and a season of peace  
And a time to be grateful, for all that we share and believe  
Let us spread love, and kindness all around  
And make the world brighter, with love that we've found  
It's the most wonderful time of the year  
With the laughter and cheer, ringing in our ears  
It's a season of love, and a season of peace  
And a time to be grateful, for all that we share and believe  
So let's embrace the season, with open hearts and smiles  
And make this Christmas, the best one in miles!



# Result

## Topic 0: Love and Affection

### Love is all that's needed:

In a world that's so complex, and filled with pain  
We search for something, that can ease our strain  
And as we journey through life, with every step we take  
We find that the answer, is love that we make  
*All we need is love*, to light up the way  
With a heart that's so pure, we'll find peace each day  
*With love in our hearts*, we'll conquer all fear  
And we'll know that in this world, love is all we need here  
With every smile, and every gentle touch  
*Love has a way, of healing us so much*  
*And as we give love*, it returns tenfold  
*With love in our lives*, we'll never grow old  
All we need is love, to light up the way  
With a heart that's so pure, we'll find peace each day  
With love in our hearts, we'll conquer all fear  
And we'll know that in this world, *love is all we need here*  
*Let us spread love*, and kindness all around  
And make the world brighter, *with love that we've found*  
All we need is love, to light up the way  
With a heart that's so pure, we'll find peace each day  
With love in our hearts, we'll conquer all fear  
And we'll know that in this world, love is all we need here  
So let's embrace love, with open hearts and smiles  
And know that in this world, *love is all we need to find!*



# Result

## Topic 1: Actions and Expressions

### A less yellow submarine:

Sailing under the sea, in a metal machine  
Exploring the unknown, in the deep blue marine  
Through the currents and waves, **we journey on**  
In this submarine, **we'll never** be alone  
**Diving** in the ocean, **exploring** all we can see  
With the submarines, **we'll conquer** the deep blue sea  
With every dive, **we discover** new mysteries  
In the underwater world, **we'll make** our own histories  
Gliding through the depths, with lights shining bright  
With every mile we travel, **we're venturing** into the night  
Watching schools of fish, and creatures of the sea  
In this metal vessel, **we're living** so carefree  
Diving in the ocean, exploring all we can see  
With the submarines, we'll conquer the deep blue sea  
With every dive, **we discover** new mysteries  
In the underwater world, we'll make our own histories  
Through the endless ocean, we'll journey on and on  
With submarines, we'll never be gone  
Diving in the ocean, exploring all we can see  
With the submarines, we'll conquer the deep blue sea  
With every dive, we discover new mysteries  
In the underwater world, we'll make our own histories  
And when the mission's done, and **we return to shore**  
**We'll remember** all the wonders, that the ocean has in store.



# Limitations

- Fixed number of topics
- Bag of words
- Lack of Sufficient Data
- Limited Vocabulary
- Unstable Results
- Difficulty in Interpreting Topics



## Future Work

- Hyperparameter tuning
- Exploring alternative topic modeling techniques that may be more suitable for small datasets
- Combine LDA with other clustering methods

---

# Q&A

---

**Thank you for listening!**