

Applied Data Science Project1 zh2404

Hip-Hop Lyrics Generator



Nowadays, there are many lyrics generators online. You can input any word you want and get the lyrics in the blink of an eye. The purpose of this data story is try to mimic online lyrics generator to some extent.

The reason why I choose hip-hop, this specific type of music, is that lyrics are much more important to hip-hop than any other. The core of hip-hop is communication. Rappers express their attitude towards life through the lyrics they write. And it is very common that rappers may often use the same beat (especially in freestyle battle), but never share the same lyrics.

Another personal reason is that my favorite artist is a rapper. I just want to test, in a data-driven perspective, can I write hip-hop lyrics like him in somehow?

General Image of Hip-Hop Lyrics

First, I can get a general image of hip-hop lyrics.

From the wordcloud, the most frequently used words are “love”, “shit”, “girl”, “baby” and so on.

The results show that most of the hip-hop songs express affection, and **love & peace** is always the leading theme of hip-hop. At the same time, because of the nature of hip-hop music, there are many **informal words** as well as **abbreviations**.



Choose A Topic

For different topics, the selected words are also different, so I conduct topic modeling analysis on the data.

Here, I use **LDA Topic Modeling** to analyze lyrics and find out the hidden topics in hip-hop. After some trials, I find out there are 6 topics in hip-hop music as follows:

Topic 1: Girl & Love

Topic 2: Depression

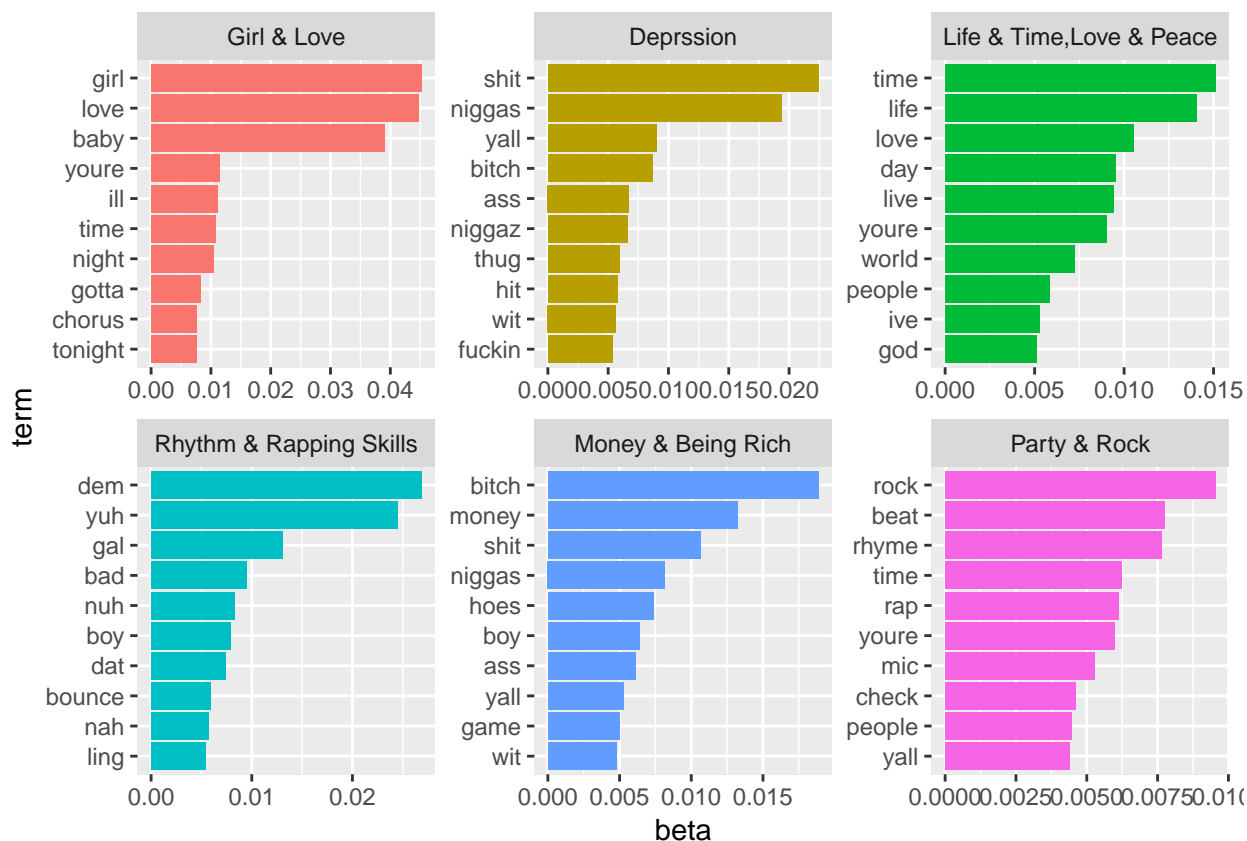
Topic 3: Life & Time, Love & Peace

Topic 4: Rhythm & Rapping Skills

Topic 5: Money & Being Rich

Topic 6: Party & Rock

For each topic, I choose 10 of most frequently used words, then I can get the histogram as follows:



Then I extract some lyrics from each topic as follows:

Topic 1: Girl & Love

"I'm sitting here lookin at your body
Baby baby baby
I've never seen something so beautiful"

Topic 2: Depression

"Now is you bout it or just talking if so then you'se a bitch
Got to get up and make it hanppen if not then you'se a bitch"

Topic 3: Life & Time, Love & Peace

"Trusted in a world I don't know"

Topic 4: Rhythm & Rapping Skills

"Oh, hey, uh, uh, uh, uh
Get mad gal, lose yuh face gal
Lose yuh head gal, nah nuh head gal
Wine yuh waist gal, lose yuh face gal"

Topic 5: Money & Being Rich

"Get the money man, the root of all evil"

Topic 6: Party & Rock

"Party people
Party people
Can y'all get funky?"

It turns out that these lyrics match their topics very well. However, there are cases that one lyric may contain several topics. For example, I find out that the lyrics in topic 6 also appears in topic 4. This means that some rappers don't just use a single topic to create their lyrics.

Emotion Type

Then I need to choose a emotion type of my lyrics.

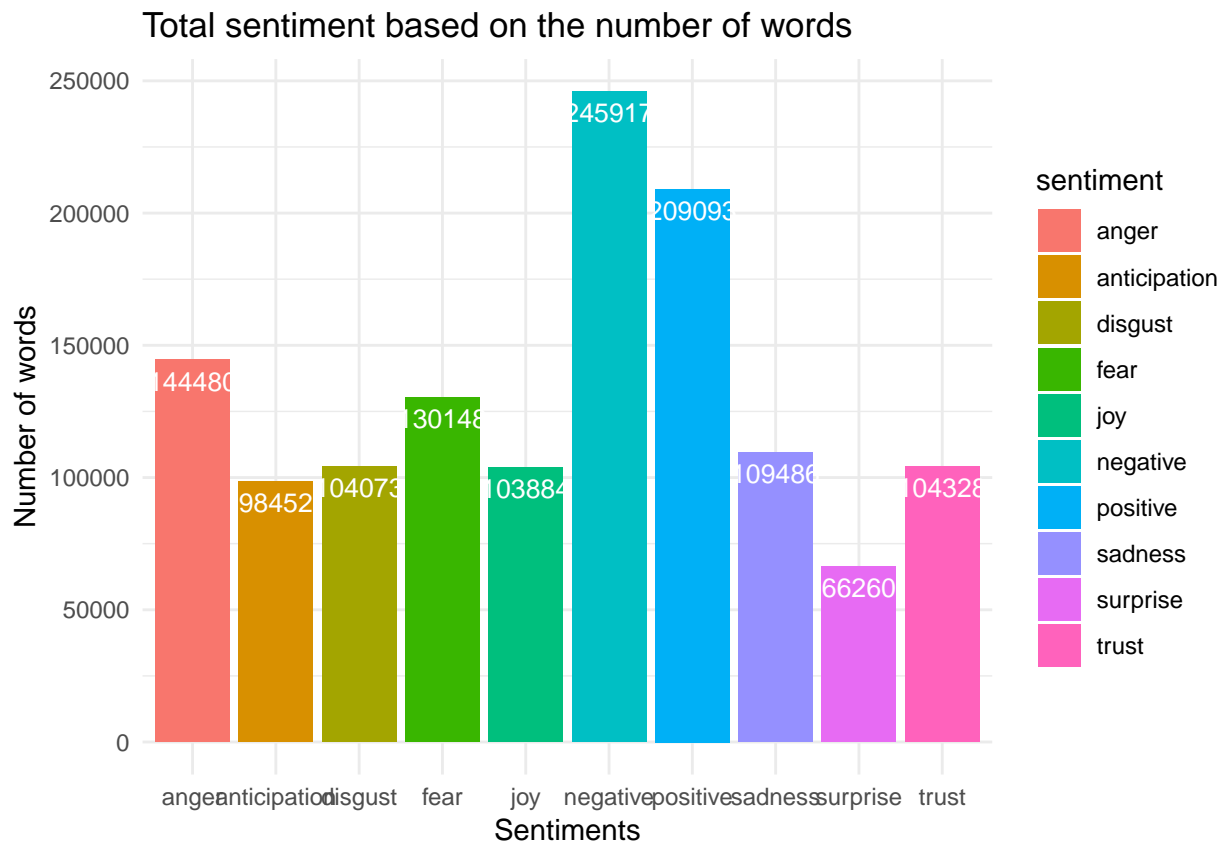
Which one is more preferable?

A broken-heart love story or a great happy ending?

So I conduct the **Sentiment Analysis**, and there are 10 emotions which are **anger**, **anticipation**, **disgust**, **fear**, **joy**, **sadness**, **surprise**, **trust**, **negative** and **positive**.

The histogram shows the total number of words in each emotion.

The results show that among all hip hop songs, the number of words expressing negative emotions is the most, while the number of words expressing surprise is the least.



Then I assign anger,disgust,fear,sadness and negative as a general negative emotion, and others as a general positive emotion.

It turns out the ratio of negative emotion to positive emotion is **1.2**, which means that the words that express the two emotions in hip-hop songs are not much different in numbers, but the negative emotions are slightly more. From this, I guess that most rappers may tend to express their dissatisfaction rather than their happy moments in their songs.

```
## [1] 1.26131
```

Lyrics Length

How many words does a hip-hop song contain in general?

At last, I analyze the length of hip-hop lyrics. As the results shown below, in average, a hip-hop song contains 196 words. And one hip-hop song can contain at most 1711 words and at least 2 words.

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	2.0	124.0	190.0	195.8	255.0	1711.0

Let's Generate Hip-Hop Lyrics

Choose several topics to write lyrics

Based on the analysis above, I plan to choose **topic 1, 4 and 6** to write a **not “that negative”** song which contains **196 words**.

According to the statistics, I compute the ratio of each topic, which is the number of words from each topic to the total number of words, and take each ratio as the probability of each topic in a song.

##						
##	1	2	3	4	5	6
##	1576	1793	2142	285	1735	1380

Get the number of words of each topic in my lyrics

According to the ratios I computed above, I can get the number of words in each topic. So there will be 95 words from topic 1, 17 words from topic 4 and 84 words from topic 6 in my song.

Then I pick up 95 of the most commonly used words in topic 1, 17 of most commonly used words in topic 4 and 84 of most commonly used words in topic 6 to form my hip-hop lyric.

The following is 30 of the most frequently used words in every topic.

##	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6
##	[1,] "girl"	"shit"	"time"	"dem"	"bitch"	"rock"
##	[2,] "love"	"niggas"	"life"	"yuh"	"money"	"beat"
##	[3,] "baby"	"yall"	"love"	"gal"	"shit"	"rhyme"
##	[4,] "youre"	"bitch"	"day"	"bad"	"niggas"	"time"
##	[5,] "ill"	"ass"	"live"	"nuh"	"hoes"	"rap"
##	[6,] "time"	"niggaz"	"youre"	"boy"	"boy"	"youre"
##	[7,] "night"	"thug"	"world"	"dat"	"ass"	"mic"
##	[8,] "gotta"	"hit"	"people"	"bounce"	"yall"	"check"
##	[9,] "chorus"	"wit"	"ive"	"nah"	"game"	"people"
##	[10,] "tonight"	"fuckin"	"god"	"ling"	"wit"	"yall"
##	[11,] "call"	"motherfucker"	"gotta"	"inna"	"white"	"hit"
##	[12,] "body"	"bone"	"mind"	"chorus"	"hit"	"ill"
##	[13,] "shake"	"gun"	"ill"	"wine"	"bout"	"black"
##	[14,] "gon"	"run"	"chorus"	"dance"	"talk"	"sound"
##	[15,] "shes"	"chorus"	"die"	"run"	"gotta"	"style"
##	[16,] "waiting"	"dog"	"heart"	"watch"	"gon"	"stop"
##	[17,] "stop"	"real"	"stay"	"stop"	"call"	"boom"
##	[18,] "leave"	"street"	"change"	"pon"	"real"	"play"
##	[19,] "boy"	"ill"	"eyes"	"mek"	"hood"	"dance"
##	[20,] "party"	"gon"	"didnt"	"till"	"chorus"	"words"
##	[21,] "mind"	"fucking"	"soul"	"walk"	"time"	"music"
##	[22,] "lady"	"kill"	"talk"	"wid"	"car"	"kick"
##	[23,] "hand"	"shot"	"friend"	"dis"	"verse"	"flow"
##	[24,] "hold"	"die"	"hes"	"girl"	"dope"	"party"
##	[25,] "imma"	"head"	"hate"	"solo"	"smoke"	"shit"
##	[26,] "play"	"stop"	"hard"	"call"	"ride"	"brother"
##	[27,] "feeling"	"time"	"wrong"	"start"	"street"	"call"

```
## [28,] "talk"      "ima"          "home"  "bye"  "move"  "step"
## [29,] "bout"      "bust"         "dream"  "cha"  "cash"  "start"
## [30,] "heart"     "ready"        "call"   "time"  "hook"  "rappers"
```

Now I can write my own lyrics!

Note that: There are 174 unique words in total. In order to make each sentence clear and coherent, and rhyme sentences as much as possible, I add and drop some words appropriately and some words I used more than once.

Finally I write a song contains 172 words in total.

Demo

Girl you know I love you so much
 Gotta go here you shake your body
 Dance like a roll I party with my bro
 Now I'm feeling hot then we take it slow
 I've got my new watch and damn it's fake
 Riding on my Porsche
 Then we talk about the story
 Oh that girl
 She wanna some fun
 You know the beat
 Watch my flow ladies love my biz
 I don't have too much time to show you my rhyme
 But frankly bro I'm not a Stan wow
 It's cool it's time for me to move
 Just tell me how you do it but you hide the truth
 Check my crew we have a new shit
 Kick your ass baby you don't know me
 Crossing the crowd and I need some money
 But I see her in the movie hunted by a zombie
 Damn it's fucking hard
 Every step I take is just another start
 I'm nobody just rotten in money of mine
 Now I just need a bottle of wine

Conclusion

1.Although I call this data analysis process a lyrics generator, it is not a generator in the strict sense. If all the words are randomly arranged and combined, the final lyrics may be incoherent and the message conveyed is a little bit weird. Therefore, in the end, I arranged all the words artificially to form lyrics.

2.Following the steps,finally I get the lyrics. However, when we look closely at the lyrics, we can see that the context may not be coherent, and more importantly, the entire lyrics may only convey some **trivial information**. It does not convey any complex feelings or thoughts.

Overall,indeed,we can generate hip-hop lyrics in this way. But it needs **life experience, professional skills** and **passion for hip-hop music** to write great hip-hop lyrics.



Figure 1: Tape in fire

Refernece

- 1.Preetish.(2018,July 18).Data Visualization and Analysis of Taylor Swift’s Song Lyrics.*Promptcloud*.Retrieved Feburary 1,2020,from <https://www.promptcloud.com/blog/data-visualization-text-mining-taylor-swift-song-lyrics/>
- 2.kratiSaxena.(2018,Febrary).Let’s analyze the song lyrics!*Kaggle*.Retrieved Feburary 2,2020,from <https://www.kaggle.com/kratisaxena/let-s-analyze-the-song-lyrics?scriptVersionId=4468644>
- 3.Bence Tóth.(2019,May 18).From Metallica to Adele — Text Analysis of successful song lyrics with R.*towardsdatascience*.Retrieved January 30,2020,from <https://towardsdatascience.com/text-analysis-of-successful-song-lyrics-e>