

Project 1: Emotions in Song Lyrics

Na Zhuo(nz2297)

9/16/2019



Introduction

How do you feel when listening to different songs? Some songs make people excited, even can make people dance. While some songs remind people of homesick and heart-broken memories. Using the lyrics database, which is a filtered corpus of 100,000+ song lyrics from MetroLyrics, let's find out the interesting facts hidden inside song lyrics. Particularly, we are going to investigate the following problems:

1. What are most frequently words in the songs of different emotions?
2. In different emotion types, which genre is the most?
3. What are most frequently emotions in the songs of different years and what is the trend of emotion changing over time?
4. Special case study of Rock music

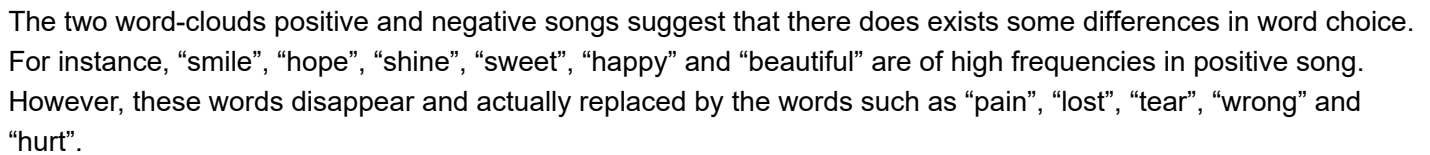
First, I import the original dataset called `df` generated by the `text-processing.rmd` file. To protect the original dataset, I duplicated it as `df.demo`, all the operations will be applied to the demo dataset. In the first piece of our data, we can find out that we have 7 variables, “song”, “years”, “artist”, “genre”, “lyrics”, “ID” and “stemmedwords”.

In this part, I conducted a sentiment analysis using *syuzhet* package. The function `get_nrc_sentiments` evaluated the stemmed word in each song and returns the positive and negative add-up score of it. I grouped the songs with higher positive score into “positive songs” group, and likewise to the negative songs.

After subsetting the data, I built the word clouds for each group.

[illegible]

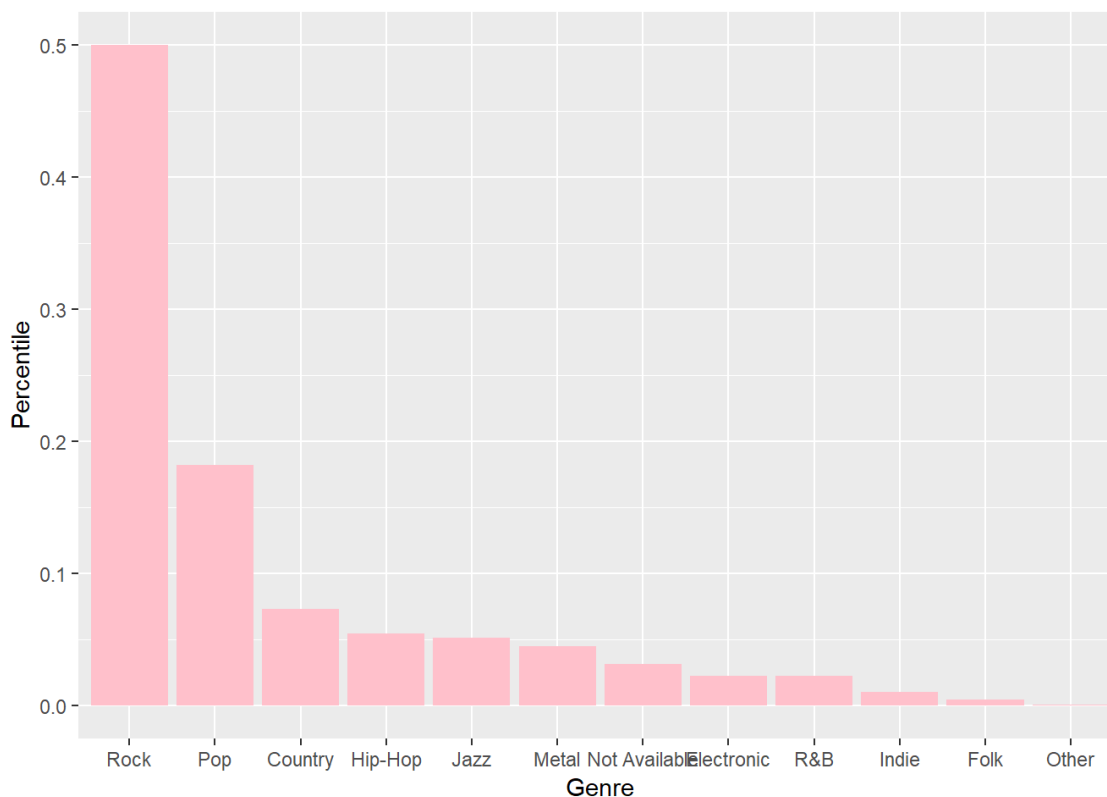
2/10



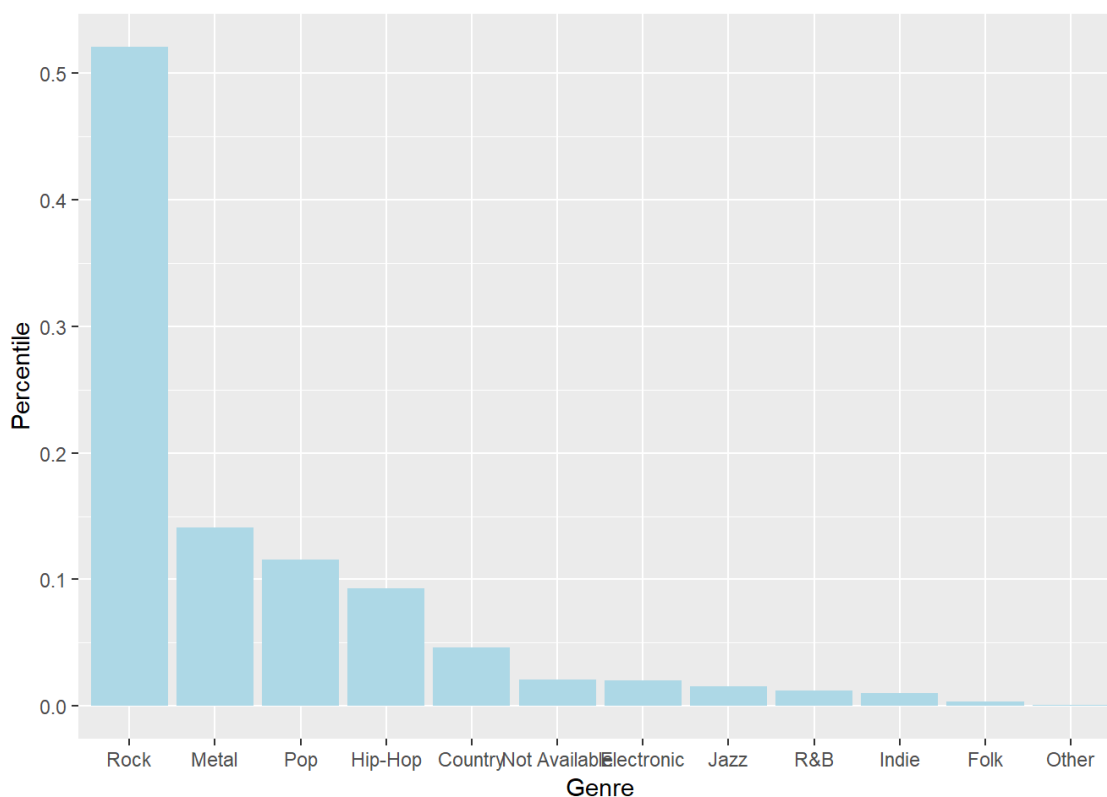
Part 2. In different emotion types, which genre is the most?

The database contains 12 types of genres, the second part of this project aims to study the majority genre of the positive and negative songs.

2.1 Positive Songs



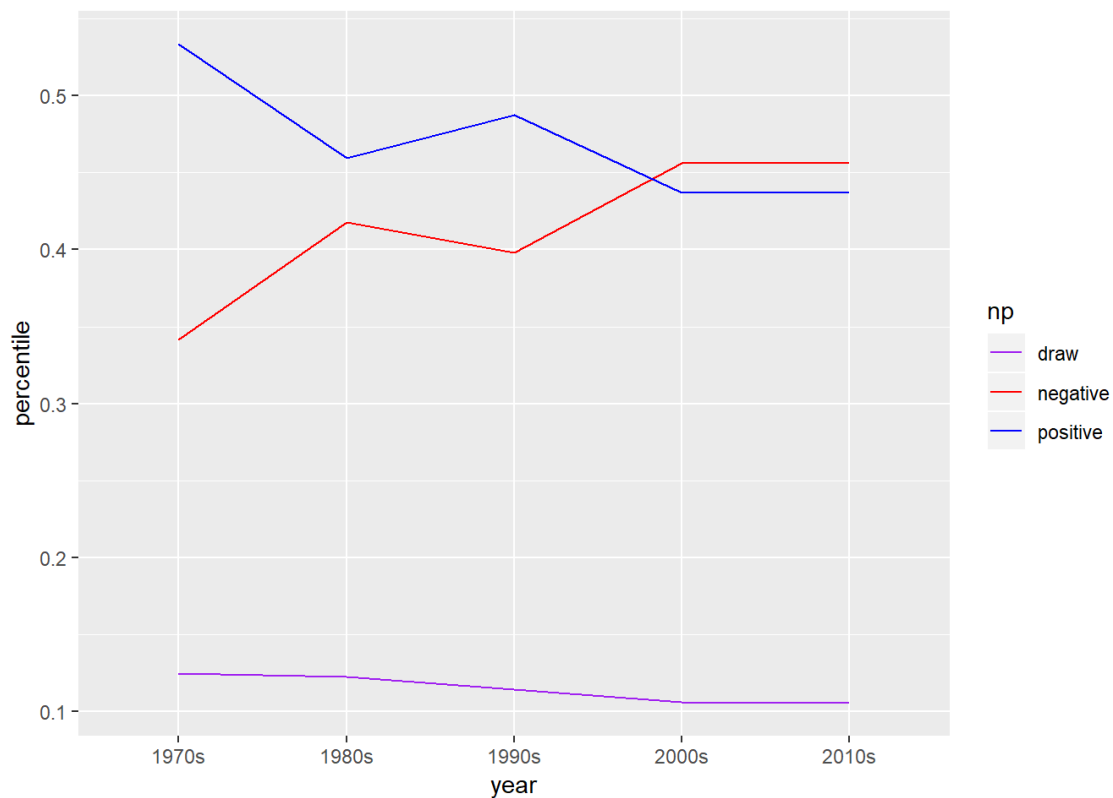
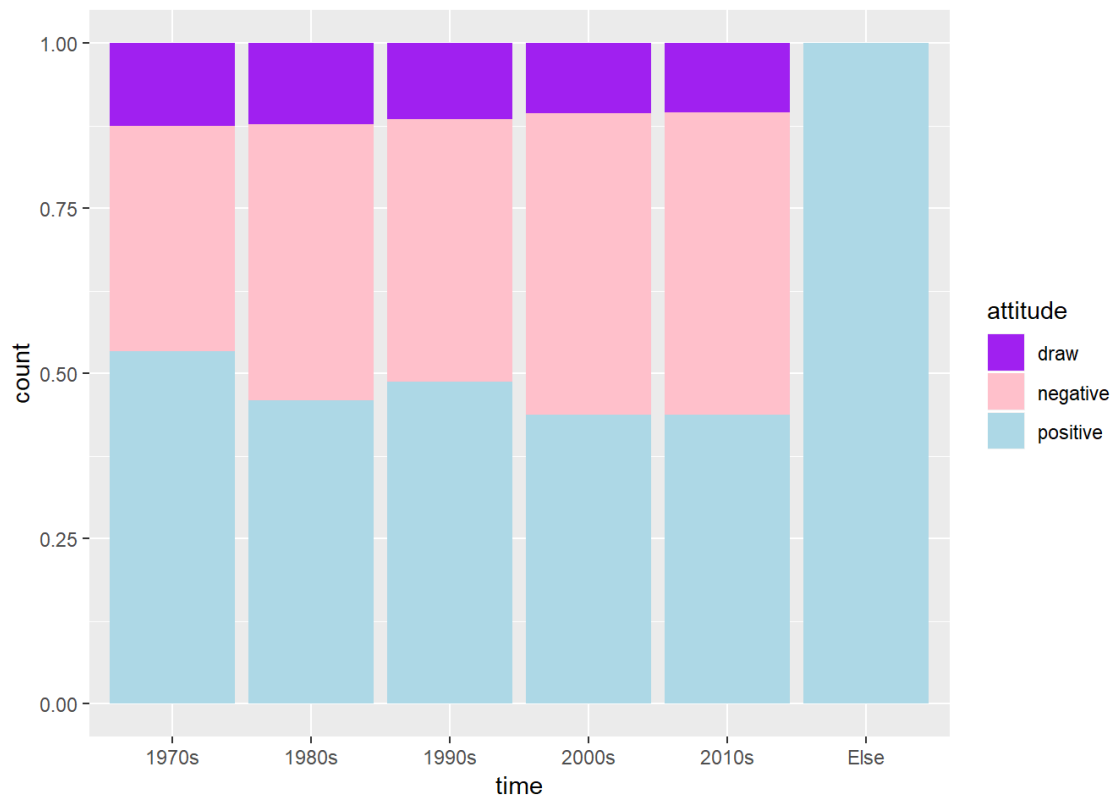
2.2 Negative Songs



The Rock music shows the first place in both group, which may due to the high weight this type of music in the original database. Except the Rock music, the most genre of positive is Pop songs and the most one of negative songs is Metal. Therefore, the music type may influence the emotion expressed through song.

Part 3: What are most frequently emotions in the songs of different years?

The third part focuses on figuring out whether there exists an emotion trend over time intervals. The years are grouped into 5 groups from 1970s to 2010s.



The figure shows the proportion of positive and negative songs in each year interval. Generally speaking according to the graph, there is some variation in proportion among year intervals.

The second graph contains the trend in proportion over time. In 1970s, there were much more positive songs than negative songs. The deviation decreased sharply in 1980s, though the positive songs remained the majority. Positive songs increased in 1990s, but not as much as 1970s. In 21th century, the negative songs took the first place until the current year.

Practically, the trend makes sense. Recently, the hottest brand-new song is Jay Chou's "Won't Cry", which caused a big crash in QQ Music when it released. It seems like nowadays we prefer to listen to songs with negative emotion.



According to the precious topics, the Rock music occupies almost half of observations. So, the behavior of Rock music has a big influence on the total trend. In the last but not the least part of this project, let's discover this type of songs with more details.

4.1.1 1970s



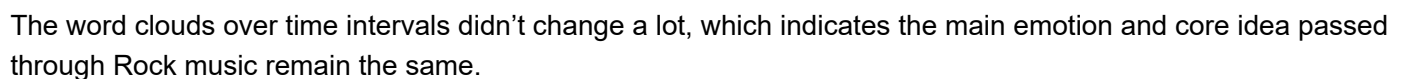
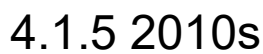
file:///C:/Users/Summer/Documents/GitHub/fall2019-proj1--nazhuo/doc/proj1_nz2297.html



4.1.3 1990s

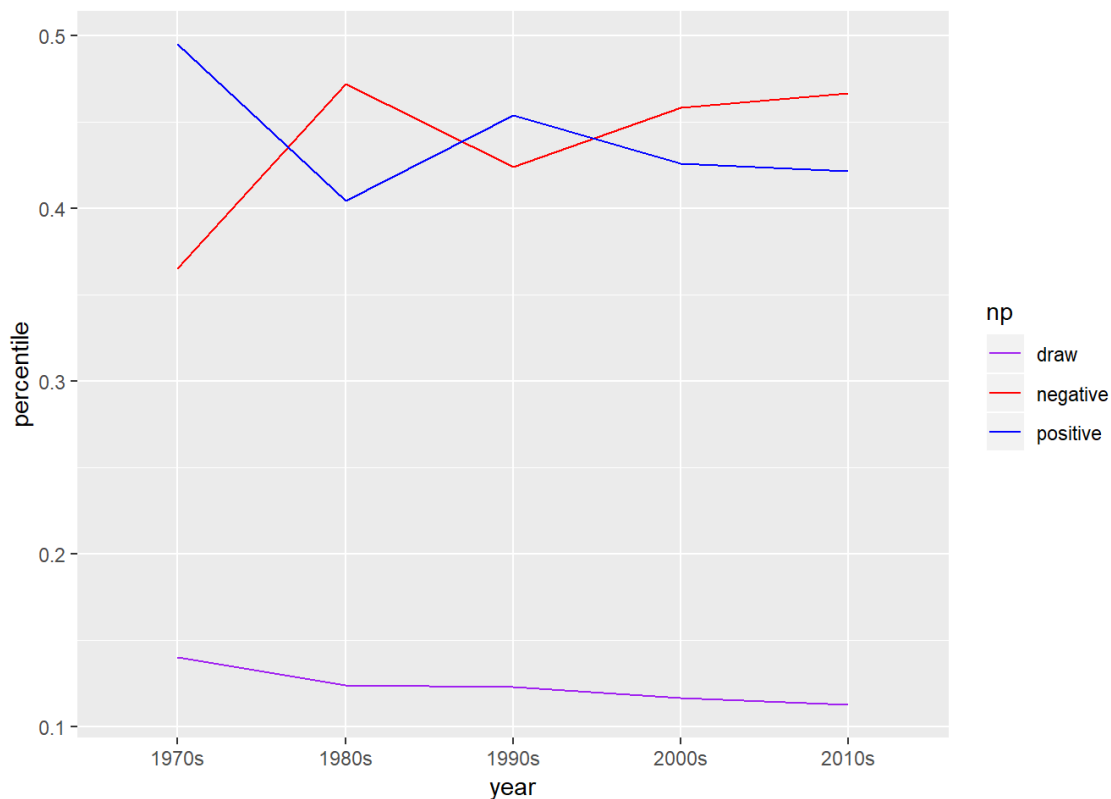


4.1.4 2000s



4.2 Emotion Trend

```
##
## 1970s 1980s 1990s 2000s 2010s Else
##    778    678    2646 50511 10231    1
```



The trend differs from the total view slightly. At the very beginning, in 1970s, there were more positive Rock songs, which is same as the main trend. However, in 1980s and 1990s, the trend is opposite to the one with all music types pooled together. In 21th century, negative music become more popular and the gap here is greater than the pooled trend.

Summary

The favorite words in lyrics differs between positive and negative songs. With Rock music remains most on both groups, Pop music tend to be more with positive emotions while metal music is more of negative emotions. Negative becomes more popular as time passing by, with positive was people's favorite last century. The keywords of a genre seem stable during time periods, but emotions expressed by lyrics varies all the time.

Refernece

- [1] Kevin Siswandi, Sentiment Analysis in R, 11, June 2017
- [2] Silge, J., & Robinson, D. (2017). Text mining with R: a tidy approach. Beijing: O'Reilly
- [3] <https://www.r-bloggers.com/sentiment-analysis-on-donald-trump-using-r-and-tableau/> (<https://www.r-bloggers.com/sentiment-analysis-on-donald-trump-using-r-and-tableau/>)