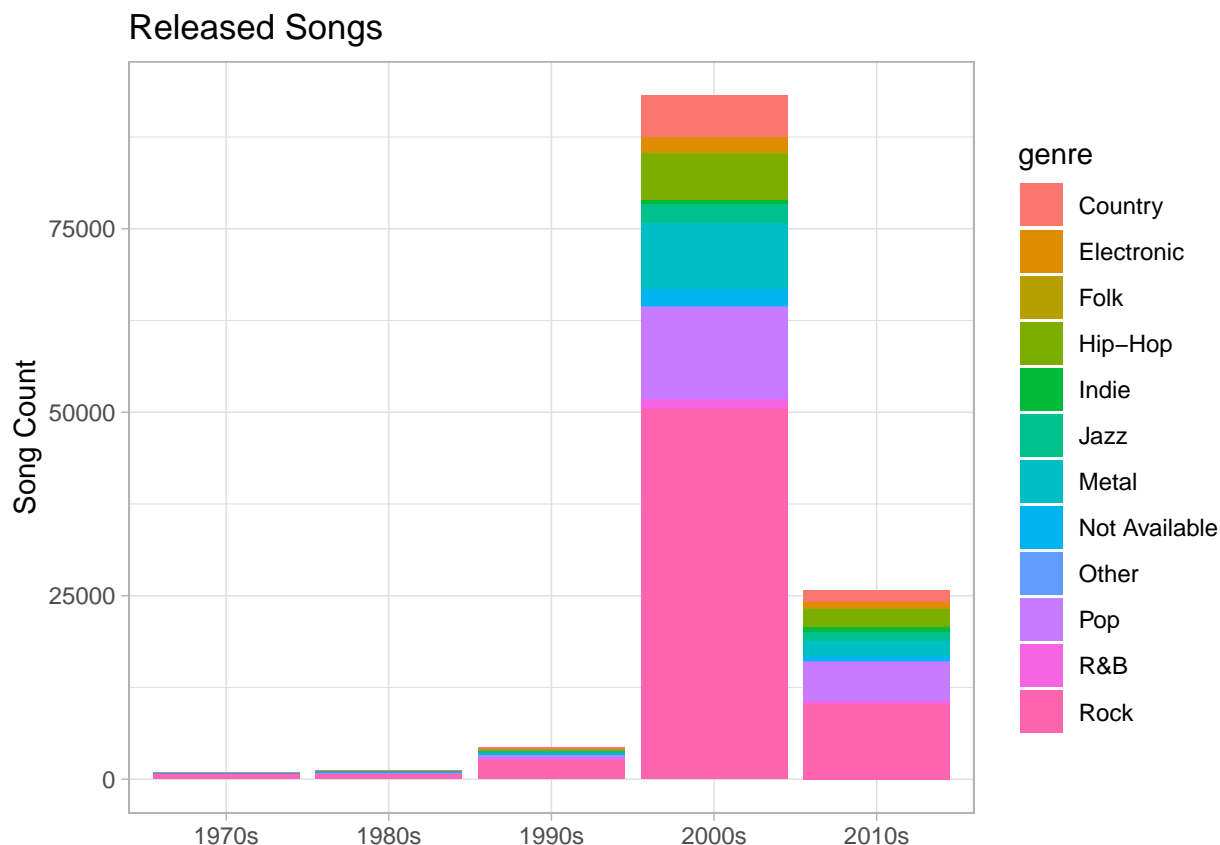


Program1: Lyrics Analysis across Time with R

Ziyang Zhang(zz2683)

Looking for song genre trends across time

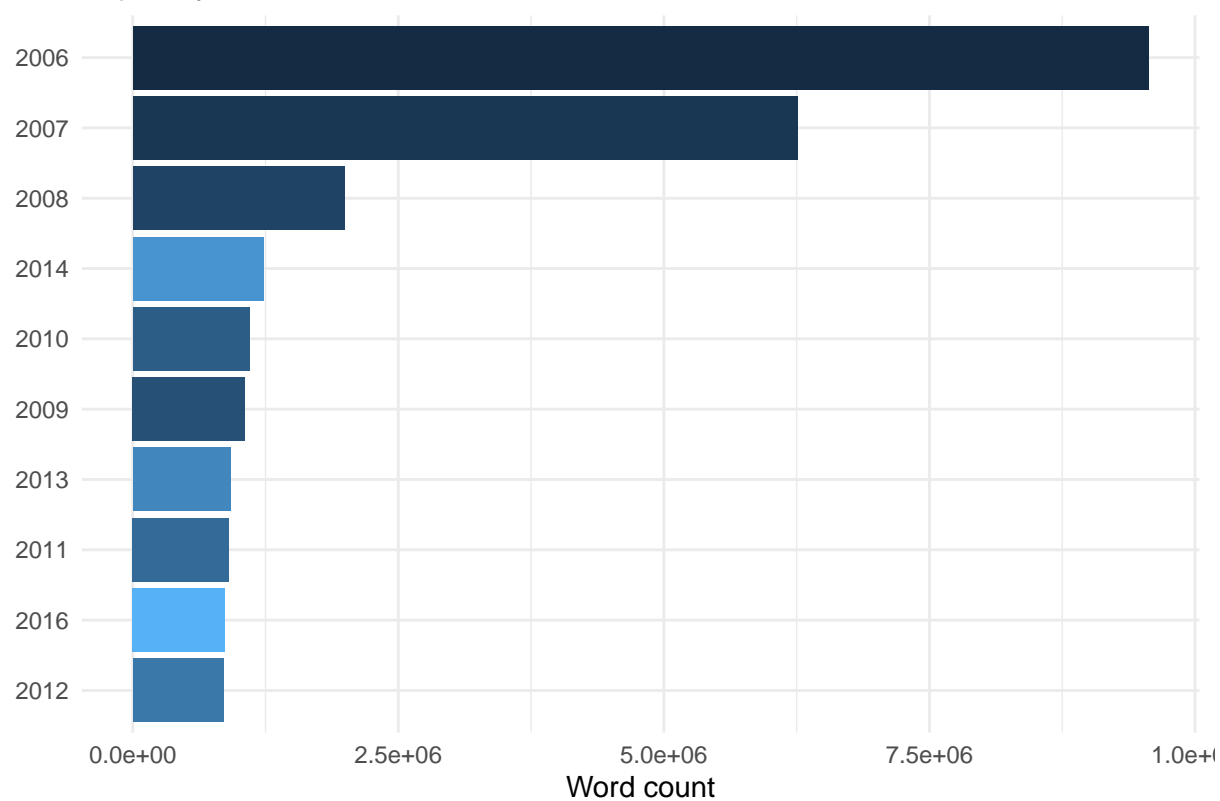
When dealing with data, because there is only one song in 1960s, we decide to ignore its effect across time(i.e. Treat this song as an outlier.). According to the histogram, we can easily find that **Rock** is always the most popular genre across time because of the large area of pink.



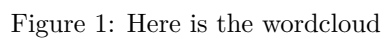
Length of songs by year

Let's find the top 10 years with the most number of words

Top 10 years in terms of word count



Let's do a simple evaluation of the most frequently used word in the full set of lyrics from 1968 to 2016.



A horizontal bar chart with a light gray background and white grid lines. The y-axis lists ten song lyrics: 'love', 'time', 'youre', 'baby', 'ill', 'day', 'life', 'heart', 'ive', and 'night'. The x-axis is labeled 'Song Count' and has major tick marks at 0, 50,000, 100,000, 150,000, and 200,000. Each lyric is represented by a solid blue horizontal bar. The bars are ordered from highest to lowest count. The 'love' bar is the longest, extending nearly to the 200,000 mark. The 'night' bar is the shortest, ending just past the 50,000 mark.

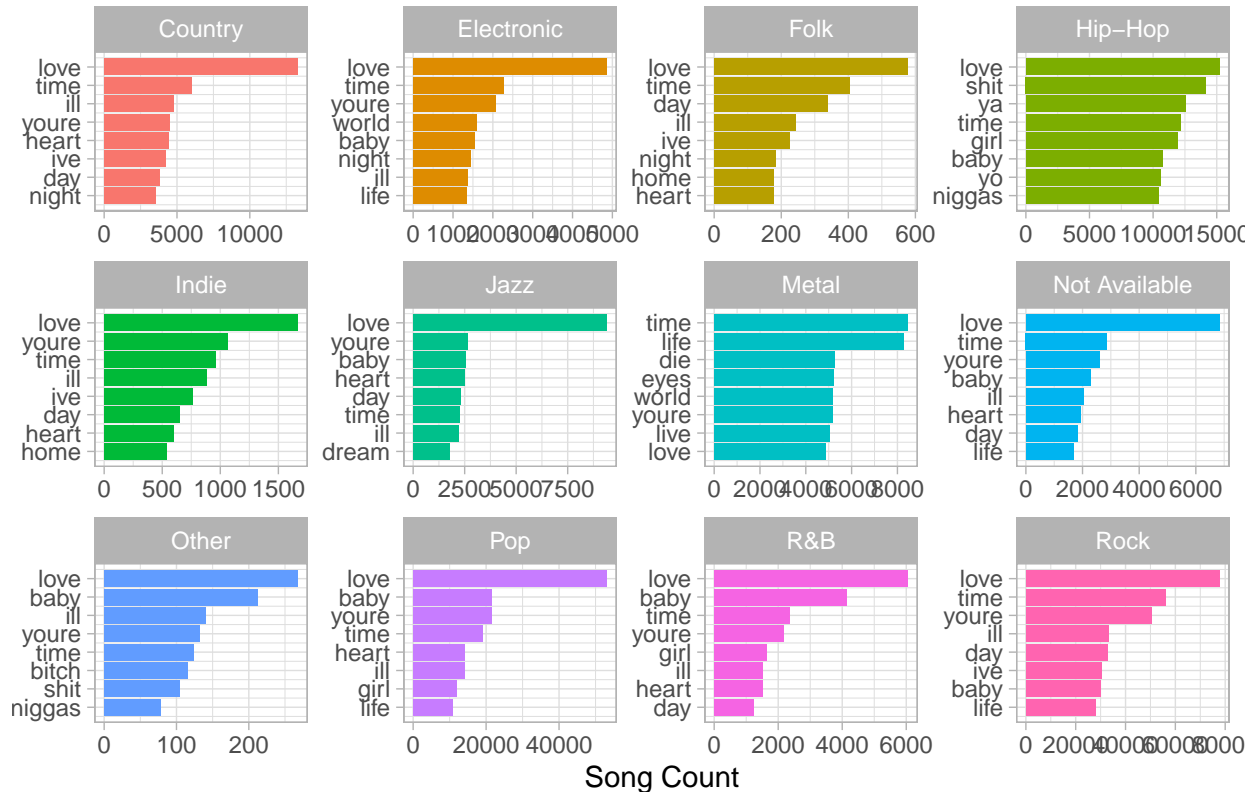
Song Lyrics	Song Count (Approximate)
love	195,000
time	115,000
youre	100,000
baby	80,000
ill	75,000
day	68,000
life	65,000
heart	60,000
ive	58,000
night	55,000

3

Popular Words in Different Genre.

So far we've just viewed the top words among all songs from 1968 to 2016. What happened if we break them up by different genre? Are some words more prevalent in songs in *Rock* verse in *Metal*?

Popular Words by Genre

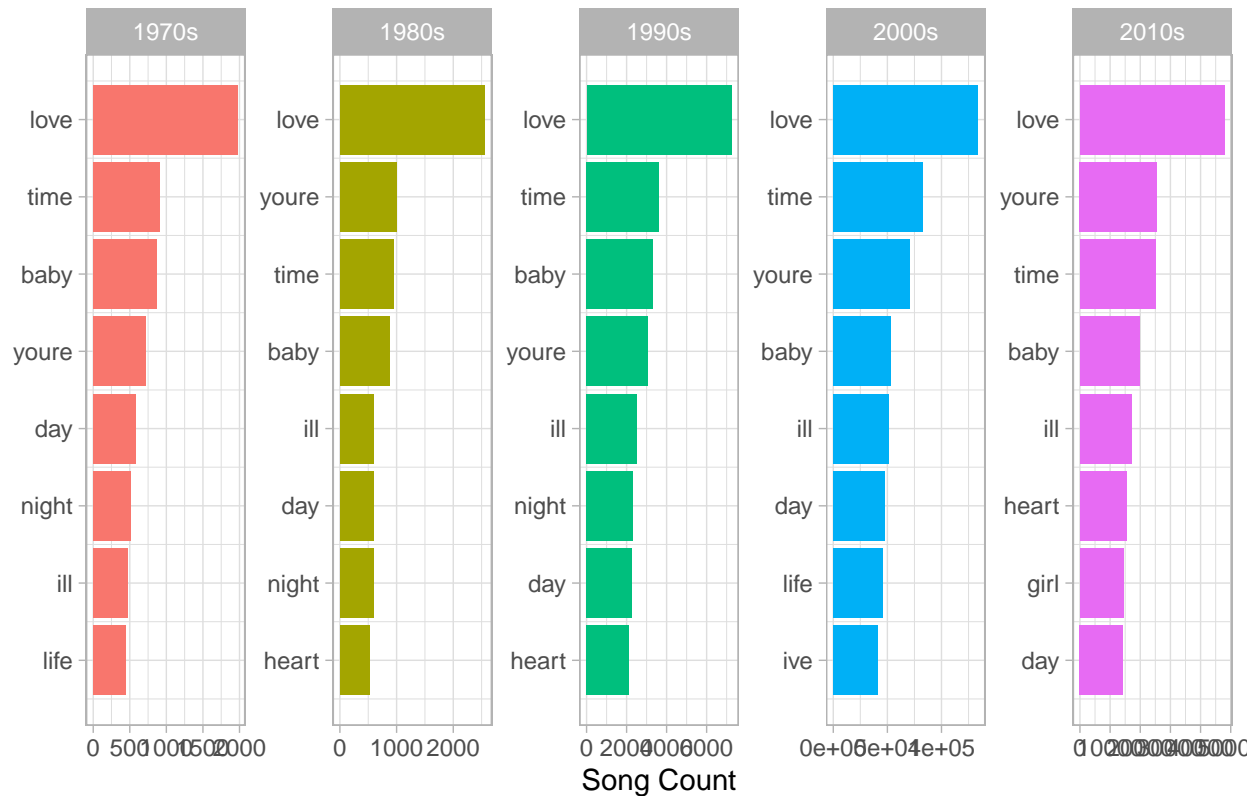


From the above, we could find that the top words across the genre levels(except Metal music) are very, very similar. Thus, it is difficult for us to determine what kind of the music based on lyrics.

Timeless Word

Some word in music are considered to be *timeless*. Timeless words perserver over time. They have passed, or have the power to pass, the test of time. They will still be enjoyed by a large number of audience. If we break down the timeline into decades, these word will roll to the top.

Timeless Words

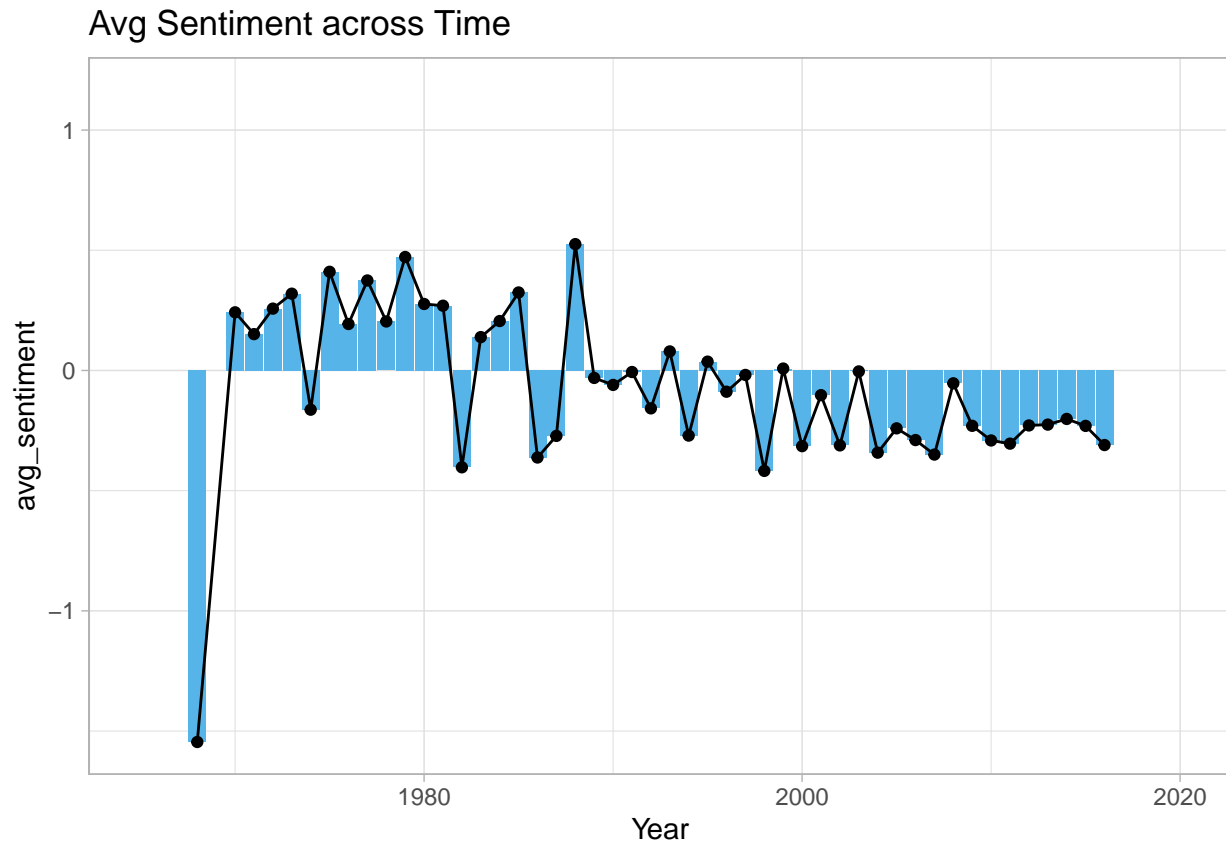


Clearly **love**, **time**, **youre**, **baby** are timeless.

Sentiment changes across time.

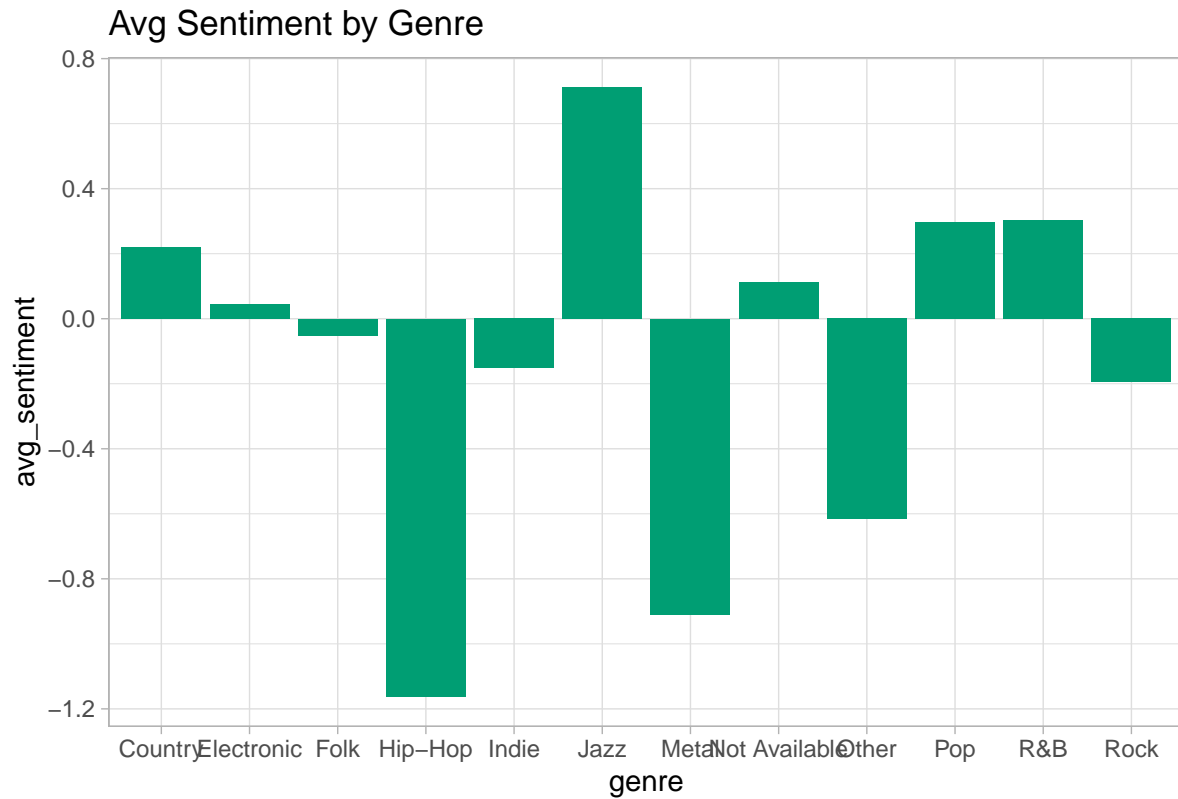
There are three sentiment dictionaries. The AFINN lexicon measures sentiment with a numeric score between -5 and 5. But the other two lexicons could only categorize words in a binary way: positive or negative. The difference between NRC and Bing is NRC has eight categories while Bing only have 2.

We want to evaluate the sentiment in a numeric way. Thus we choose AFINN lexicon to score every word in lyrics. We want to see the change of emotions from 1968 to 2016.

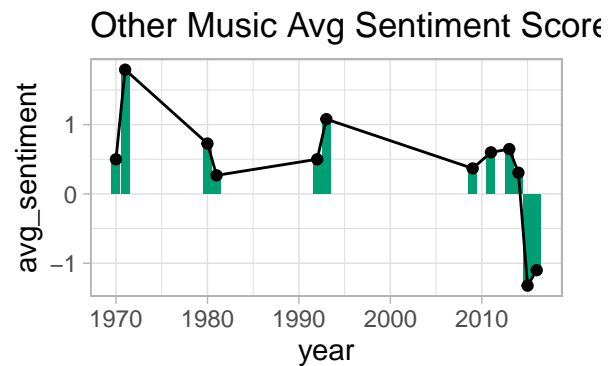
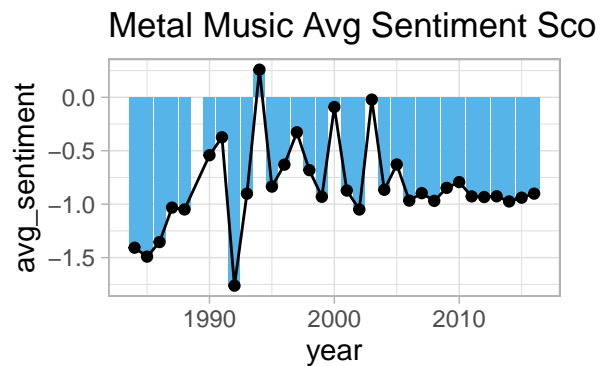
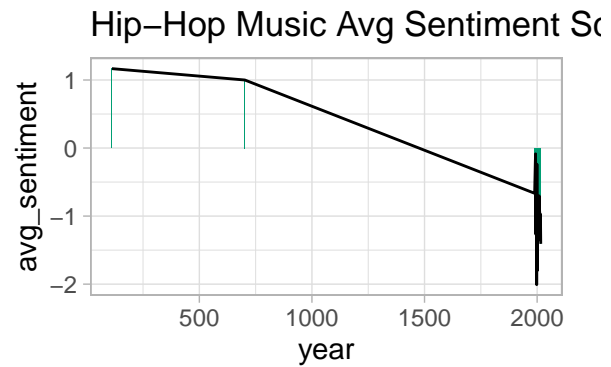
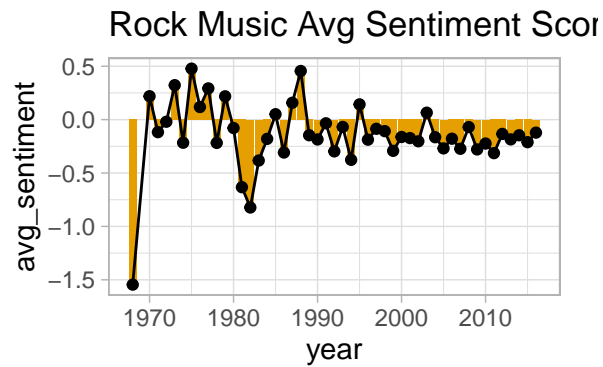


It is a little wierd that in 1968, the average sentiment score is very negative. When we check what was happened in 1968, we find that there is only one song named *Without You* in 1968, which is a Rock music about broken love. Thus, it is not hard to understand why the emotion in 1968 is extremely negative.

In addition, according to this figure, we could see that the average sentiment score is almostly positive before 1988, while the average score is always negative after 1988. Why does the emotion change from positive to negative so obviously?



From the above figure, we could find that the metal, Hip-Hop, Rock and Other music have relatively negative emotions.



From the above four plots, we can see that the Rock music emotion trend is quite similar to the whole music

trend across time. The Metal music is always negative (except in 1992) and it does not display the change from positive to negative after 1988. We could also ignore the Hip-Hop music and other music effect on the whole emotion change because of their relatively small numbers of songs.