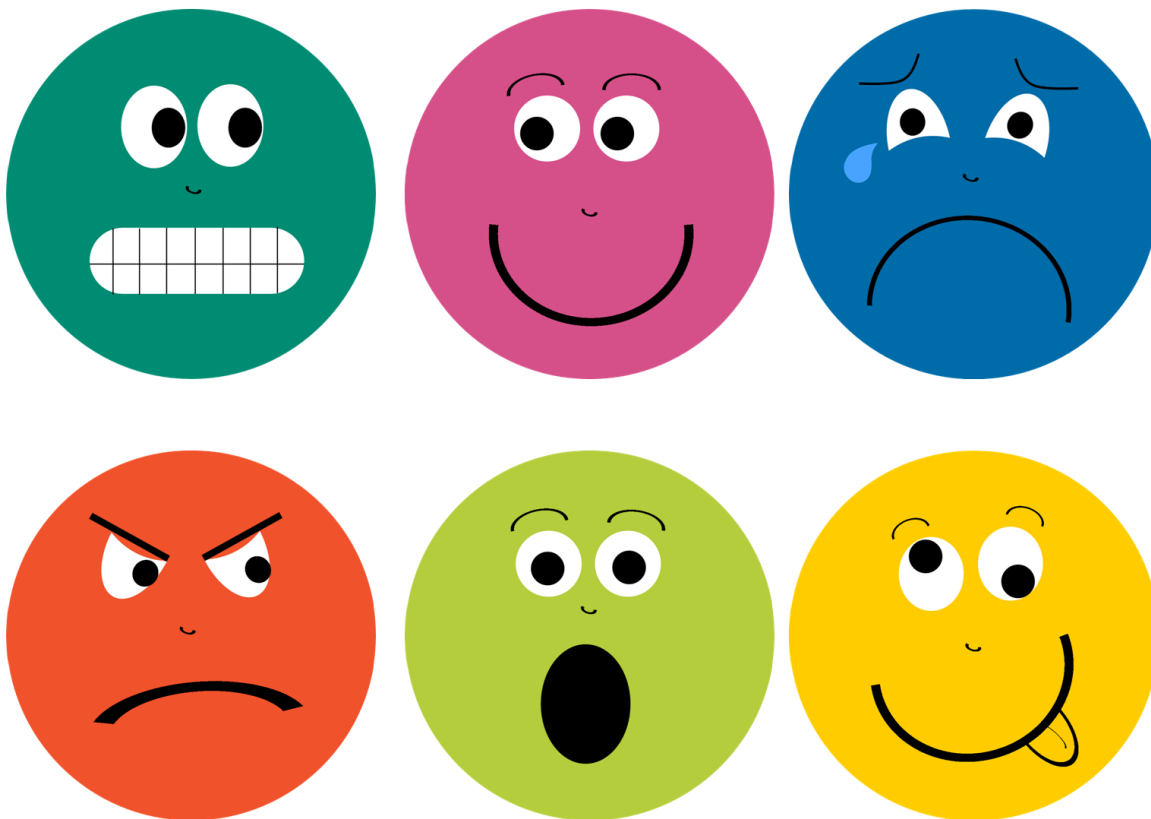


Emotion patterns in different lyrics genres



We know that lyrics will contain countless emotions, anticipation, joy, sadness and so on. But what will they differ in different lyrics genres, and also in different time period? Today, my main theme is to use sentimental analysis available in R package to explore such emotion patterns and get some interesting findings. Basically, my analysis will be conducted in three directions:

1. Compare emotions contained in words in different lyrics genres
2. Compare emotions contained in top200 most frequent words in different genres
3. Compare emotions contained in words for different time period of certain genres

First, after we import the package and filter the data we need today. We will want an overview of the data. Thus we show counts of lyrics genres in every year by a table function. Although this may not be directly related to what we discuss today. It provides us extremely clear description of lyrics genres in different years. We will then gain an overview and choose certain genres to analyze afterwards based on this chart.

##											
##		Country	Electronic	Folk	Hip-Hop	Indie	Jazz	Metal	Pop	R&B	Rock
##	2011	175	85	16	382	115	103	296	791	28	1268
##	2012	163	99	28	325	89	134	261	759	40	1354
##	2013	329	156	7	276	91	195	248	712	70	1286
##	2014	252	193	28	410	104	320	309	986	100	2058
##	2015	227	137	12	212	148	168	277	588	175	1154
##	2016	194	132	19	283	115	176	260	691	50	1237

##Compare emotions contained in words in different lyrics genres

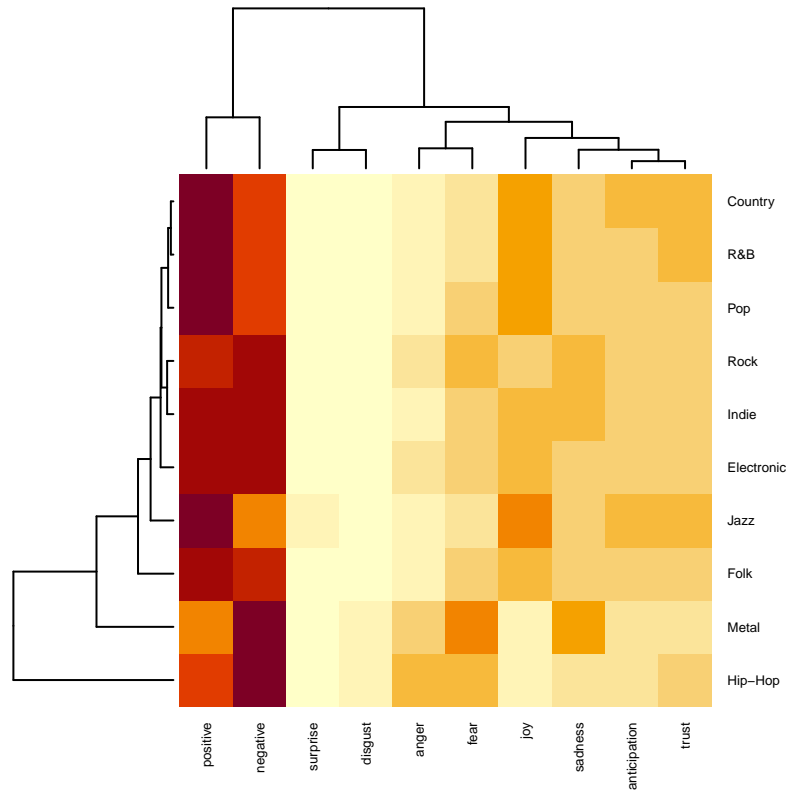
We used sentimental analysis available in R package to gain emotion scores for samples from every lyrics genres. Because the large sample space, we will sample 2000 lyrics from each genre and get means for such scores. Then what we have is the average numbers for 10 kinds of emotions in each lyrics. Besides, we would better compare emotions for same genre because average numbers of words for each genre are also different.

Let's focus on the positive percentage column which represent the percentage of positive emotions words of the genres. We can find the two most outstanding genres are Metal and Jazz, which can represent most negative and positive genres. Other genres, although have slight difference in emotions' weight, stay a pos-percentage around 50% relatively.

We also create the heatmap in which emotions are on x-axis and genres on y-axis.

##	anger	anticipation	disgust	fear	joy	sadness	surprise	trust
## Hip-Hop	9.763	7.215	6.631	9.587	6.498	7.687	5.108	7.981
## Pop	2.442	3.316	1.867	2.978	3.817	3.148	2.052	3.347
## Metal	5.378	4.265	3.870	6.983	3.484	6.189	2.565	4.178
## Rock	2.752	3.167	2.015	3.463	3.273	3.481	1.968	3.260
## Country	2.234	3.245	1.641	2.678	3.788	3.018	1.913	3.347
## Indie	2.398	3.062	1.757	3.052	3.272	3.431	1.973	3.158
## Jazz	1.692	2.908	1.287	2.046	3.834	2.556	1.775	2.784
## Electronic	2.299	2.772	1.629	2.895	2.986	2.876	1.697	2.718
## R&B	2.308	3.013	1.785	2.778	3.719	3.000	1.791	3.226
## Folk	3.108	4.118	2.279	3.839	4.397	4.094	2.330	4.151

##	negative	positive	pos_perc
## Hip-Hop	17.955	14.752	44.59684
## Pop	5.287	6.313	54.51731
## Metal	11.238	7.266	39.26303
## Rock	6.061	5.714	49.44530
## Country	4.902	5.975	55.79549
## Indie	5.472	5.450	51.21877
## Jazz	3.722	5.765	60.15721
## Electronic	5.047	5.037	50.77447
## R&B	4.922	5.904	54.40732
## Folk	6.857	7.578	52.80344



Compare emotions contained in top200 most frequent words in different genres

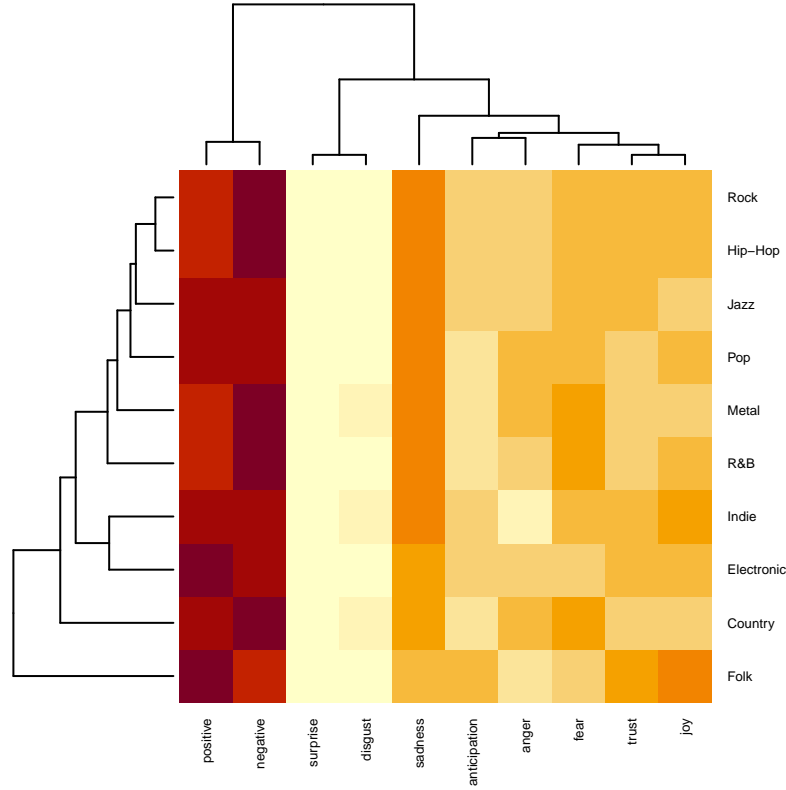
For the second part, first we separate the lyrics into words and get top200 most popular words in every genre. Then we conduct similar sentimental analysis on the wordset to see what emotions do they contain.

We find Metal still gets a low pos-percentage, which fits the above analysis. And we can confirm Metal gives more negative emotions than others.

The different finding is that we see Folk becomes relatively positive genre, which also makes sense.

However, based on the data we can see most popular words for these genres have a lot in common, so the emotion analysis also gave very similar outcomes. So we can't say this analysis became so successful except it confirms Metal's sadness pattern.

##	Hip-Hop	Pop	Metal	Rock	Country	Indie	Jazz	Electronic	R&B	Folk
## 1	love	love	love	love	love	love	love	love	love	love
## 2	time	time	time	time	time	time	time	time	time	you're
## 3	you're	you're	you're	you're	you're	you're	you're	you're	you're	baby
## 4	baby	baby	baby	baby	baby	ill	baby	baby	ill	time
## 5	ill	ill	ill	ill	day	day	ill	ill	baby	heart
## 6	day	day	day	day	ill	heart	heart	day	life	ill



##Compare emotions contained in words for different time period of certain genres

Because of Metal and Jazz's special patterns, I'll show these two genres' time analysis

We separate data into four blocks and get emotion scores for them and show the outcomes together

I also tried to conduct same method on other genres with enough data. So the final conclusion is that:

There is not an obvious change in emotions before and after 2006. But there seems to have a slight trend for positive emotions to increase as time passed.

```
##          anger anticipation  disgust    fear    joy  sadness
## emo_metal_bef 5.372954      4.207405 4.005098 6.967534 3.546820 6.253019
## emo_metal_aft 5.142096      4.074552 3.788490 6.661435 3.401532 6.011211
##          surprise    trust negative positive pos_perc
## emo_metal_bef 2.566407 4.172793 11.26643 7.248189 39.09895
## emo_metal_aft 2.478886 4.044376 10.77924 6.966368 39.29977

##          anger anticipation  disgust    fear    joy  sadness
## emo_jazz_bef 1.671419      2.918341 1.274141 2.042126 3.792612 2.491899
## emo_jazz_aft 1.635130      2.920124 1.291198 2.021714 3.854207 2.504071
##          surprise    trust negative positive pos_perc
## emo_jazz_bef 1.779650 2.802981 3.629294 5.653273 60.40425
## emo_jazz_aft 1.799147 2.853044 3.625824 5.718496 60.74849
```

##Summary

1. We got an emotion score table presenting emotions contained in samples of different lyrics genres. From percentage of positive words, we can see Metal seems to be most negative and Jazz seems to be most positive.

2. Because most popular words in these genres have many in common, we can't get an obvious outcome by analyzing these words' emotions. However, it still reinforces Metal's negative emotion and Folk become

most positive this turn.

3. We use 2006 as watershed and Metal,Jazz as samples to explore emotions vs time. We didn't find a clear trend but they show positive emotions increase very slightly by the time.