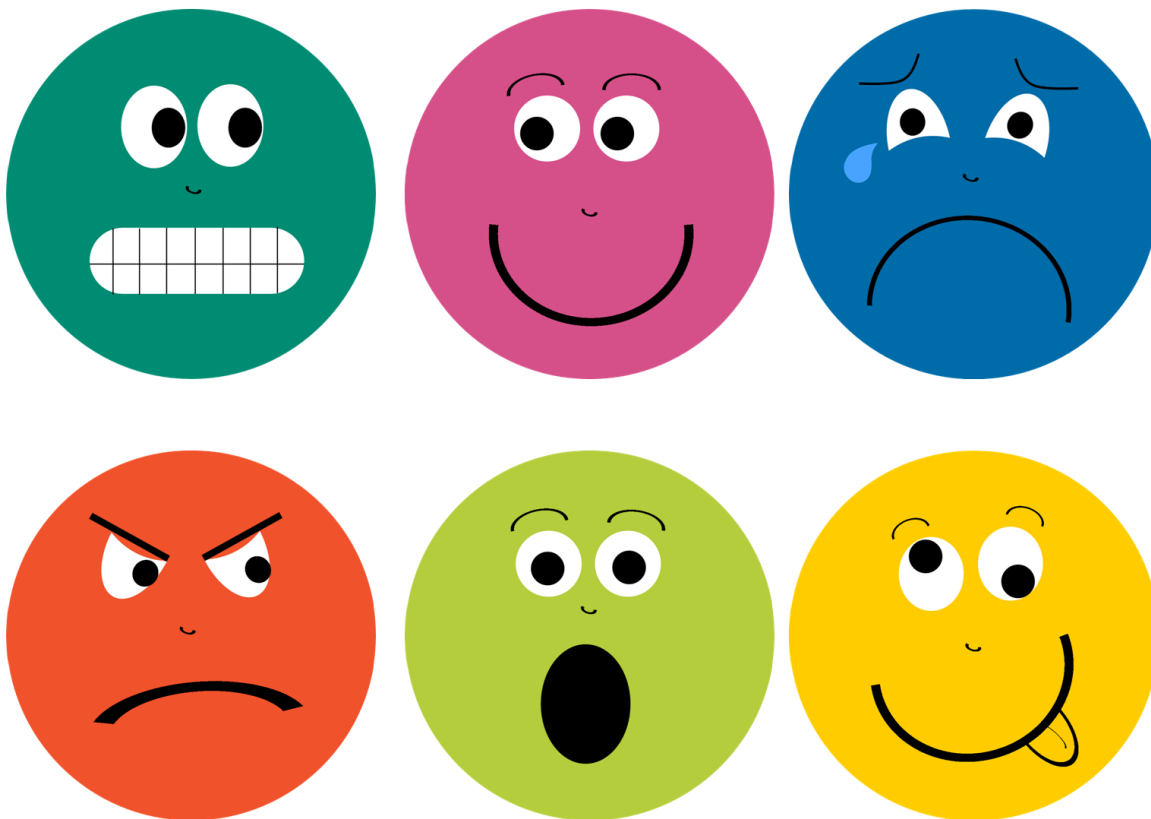# Emotion patterns in different lyrics genres

We know that lyrics will contain countless emotions, anticipatoin, 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. Basiclly, 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
```

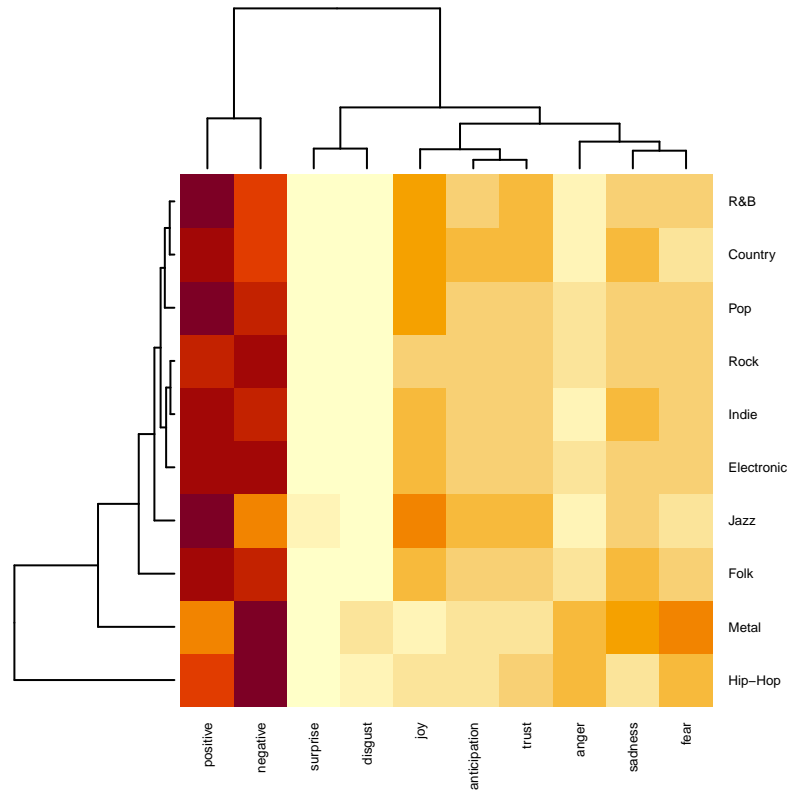##Compare emotions contained in words in different lyrics genres

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 forcs on the postive 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    9.487        6.997   6.483 9.276 6.520   7.652    4.805 7.916
## Pop        2.646        3.382   1.952 3.054 3.883   3.191    2.158 3.382
## Metal      5.311        4.101   3.877 6.839 3.400   6.021    2.541 4.259
## Rock       2.652        3.077   1.881 3.218 3.207   3.268    1.908 3.083
## Country    2.040        3.253   1.571 2.617 3.658   3.110    1.896 3.213
## Indie      2.343        3.029   1.810 3.067 3.262   3.409    1.894 3.170
## Jazz       1.669        2.903   1.284 2.069 3.807   2.609    1.810 2.781
## Electronic 2.371        2.808   1.657 2.894 2.999   2.912    1.793 2.765
## R&B        2.392        3.053   1.787 2.859 3.884   2.994    1.804 3.213
## Folk       3.169        3.884   2.395 3.900 4.101   4.106    2.262 3.997
##            negative positive pos_perc
## Hip-Hop      17.676   14.401 44.55396
## Pop           5.555    6.372 53.90583
## Metal        10.971    7.064 39.28545
## Rock          5.786    5.565 50.05201
## Country       4.838    5.745 55.61817
## Indie         5.399    5.517 51.28267
## Jazz          3.739    5.587 59.76361
## Electronic    5.047    5.130 51.01067
## R&B           4.879    6.023 54.66127
## Folk          6.972    7.142 51.00649
```

## ##Compare emotions contained in top200 most frequent words in different genres

For the second part, fisrt we seperate the lyrics into words and get top200 most popular words in every genres

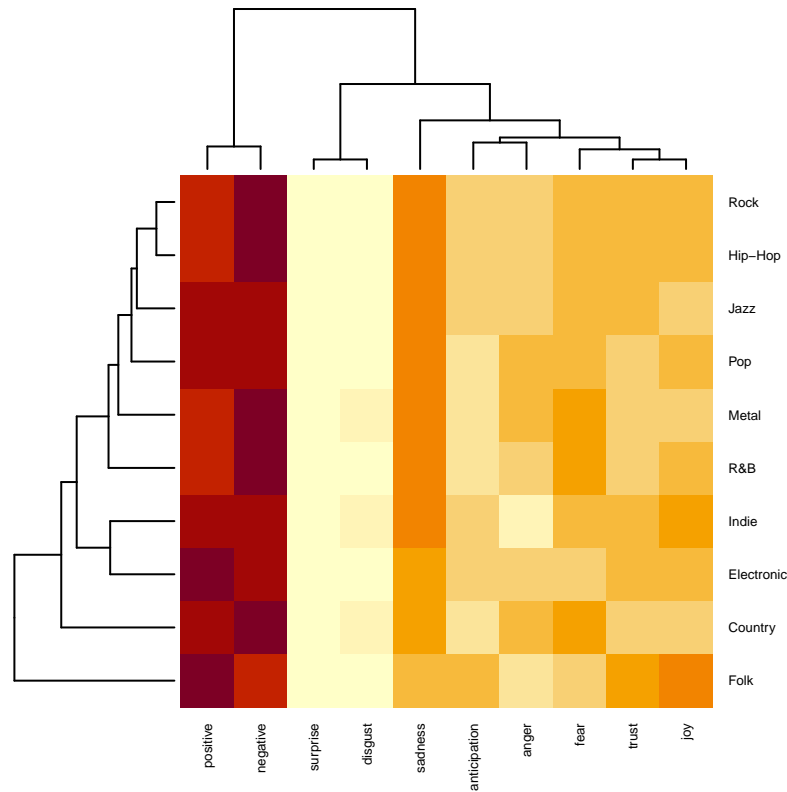Then we conduct similar sentimental analysis on the wordset to see what emotions do they contain

We find Metal still get a low pos-percentage, which fits the above analysis. And we can confirm Metal give more negative emotions than others.
The different finding is that we see folk becomes relatively positive genre, which also make 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 youre
## 3    youre youre youre youre   youre youre youre      youre youre  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

##         anger anticipation disgust fear joy sadness surprise trust
## Hip-Hop    20           19      13   22  22      25       14    21
## Pop        21           17      14   21  21      25       13    20
## Metal      22           17      16   23  20      25       14    20
## Rock       20           19      14   23  22      26       14    21
## Country    20           15      13   21  18      22       12    18
## Indie      15           20      15   22  23      26       13    22
## Jazz       20           19      14   22  20      25       12    22
```

```
## Electronic        19             19         14  20  22        23        14  22
## R&B               21             19         14  25  23        27        14  21
## Folk              15             20         11  18  24        19        12  22
##          negative positive pos_perc
## Hip-Hop         36          33 48.44444
## Pop             34          33 47.48858
## Metal           37          32 45.57522
## Rock            36          33 47.80702
## Country         33          32 46.56863
## Indie           34          34 50.00000
## Jazz            35          34 47.98206
## Electronic      32          35 50.90909
## R&B             37          34 47.23404
## Folk            29          34 54.90196
```



## 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 seperate 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.