

R Notebook

Data Story on song lyrics

Music is a great tool to tell story and express feelings we can't grasp. And music in different genres tend to convey emotion in different ways. Although the genres classifications may be subjective and even overlap, there are still some shared characteristics represented in lyrics through genre breakdown.

the distribution of each genres

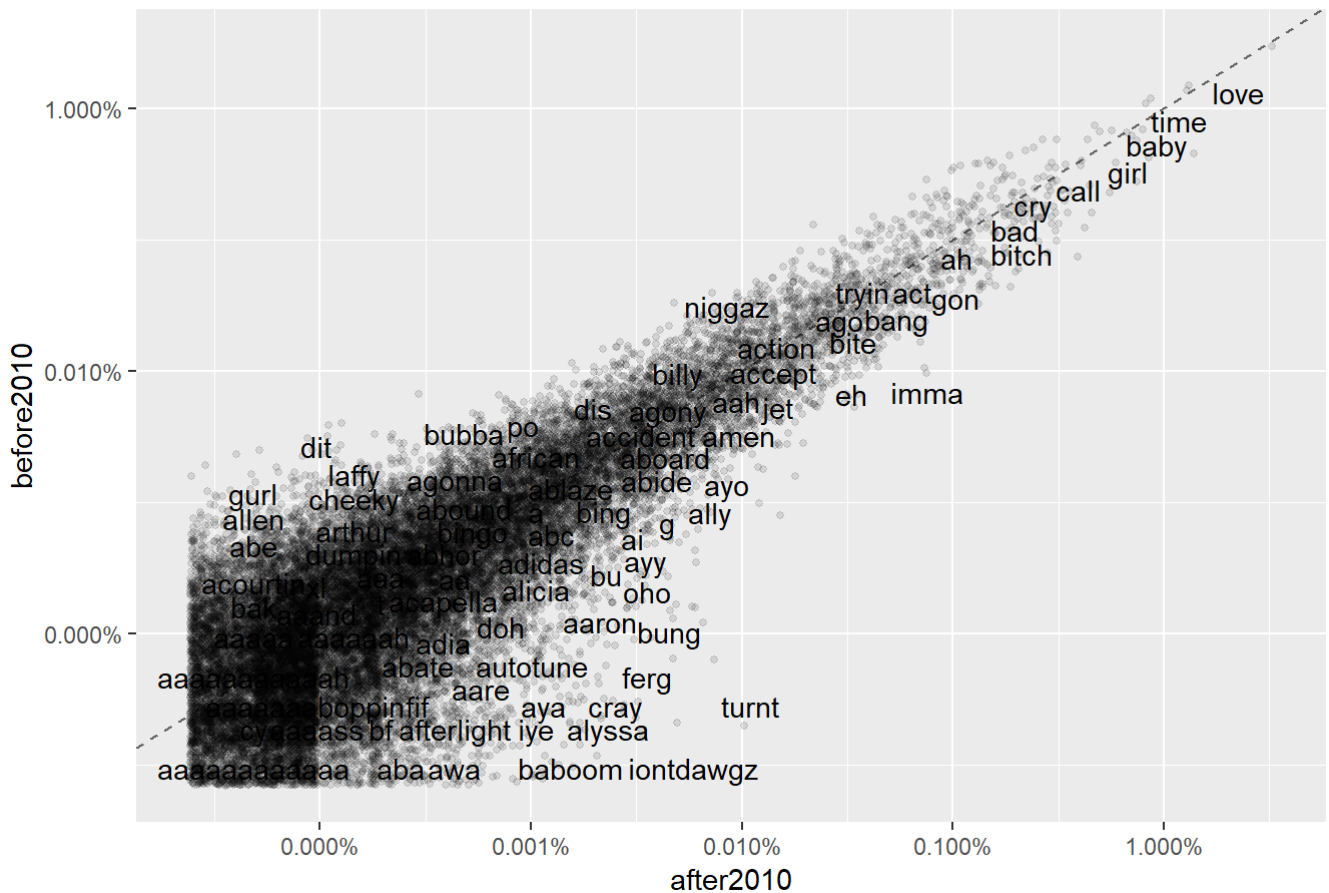
```
##
##          age1970s age1980s age1990s age2000s age2010s      Sum
## Country          49        20        219      5690      1556    7534
## Electronic        9         14         44      1843        902    2812
## Folk              10          0         27        357        120     514
## Hip-Hop           0          11        254      6268      2370    8903
## Indie             0           0          0        591        727    1318
## Jazz              80          92        186      2522      1242    4122
## Metal             0          63        184      9030      2034   11311
## Other             5          13          5          4        120     147
## Pop              163         210        478     12543      5303   18697
## R&B                65          67        179      1376        487    2174
## Rock              778         678       2646     50511     10231   64844
## Sum              1159        1168       4222     90735     25092  122376
```

According to the list, we can see the popularity of different music genres in different phases. Musicians have long-term preferences on Rock music, Pop music, and Country music, besides, Metal and Hip-hop have led a trend after 2000. folk music and Indie music seems to be the niche genre with the fewest audience.

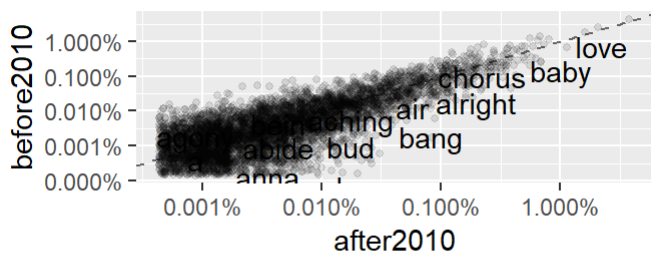
Word Analysis

Does patterns varies in song lyrics between genres and time? I try to compare the most frequently used words in different genres to get an answer. I take the 6 most popular genre—Country, Hiphop, Jazz, Mental, Pop, and Rock music— as examples, and roughly divide the time into 2 stages of before/after 2010.

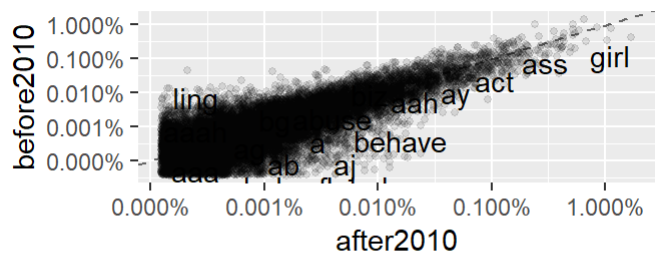
Words Proportion before/after 2010



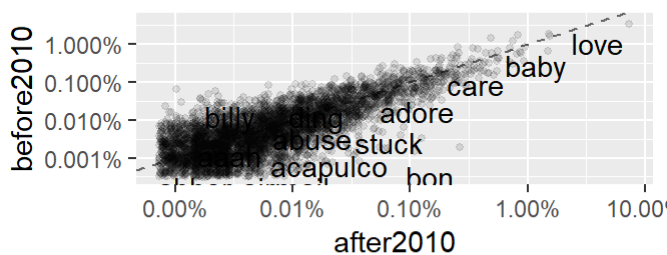
Words Proportion for Country Musi



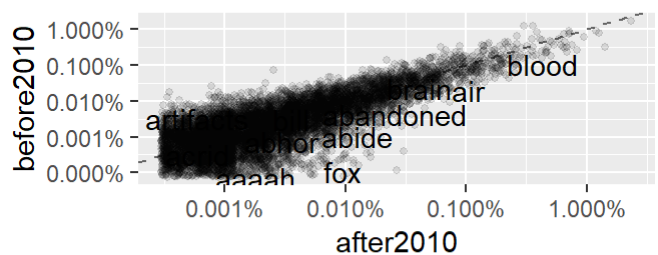
Words Proportion for Hip-Hop Mus



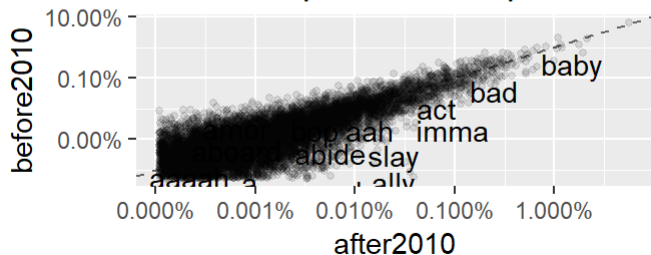
Words Proportion for Jazz Music b



Words Proportion for Mental Music



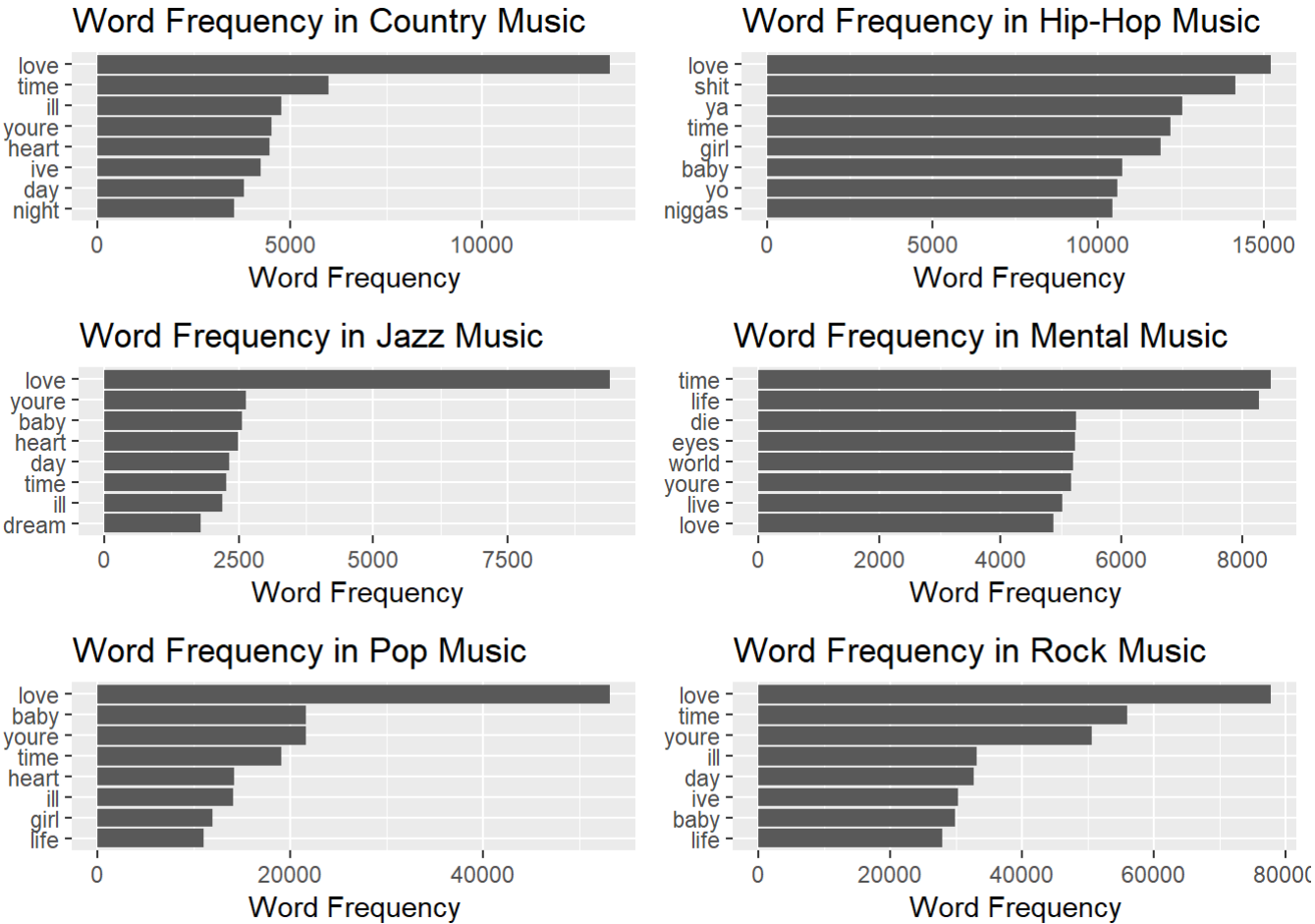
Words Proportion for Pop Music be



Words Proportion for Rock Music k



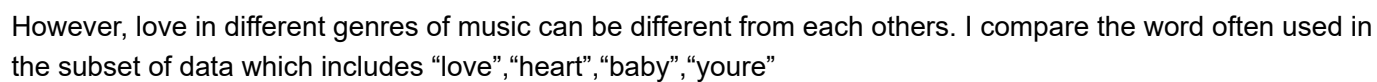
From the plots and cloud above, the word like “love”, “baby”, “girl”, “youre” are always used in lyrics all the time, words like “cowboy”, “bubba”, “artifacts” which used to frequently appear, is rarely shown any more. And words like “camouflage”, “troll”, “afterparty”, “pirate” starts to shown in lyrics in recent years.



```
## [1] 122376
```

```
## [1] 51182
```

I also list the top 8 frequently used word in every genres, which shows the similar results. among all the 122,376 lines of lyrics, 51182 of which have “love” included. The word “love” show the core emotion from old times to today in every genres, except the mental music which is presented below. Mental music tend to focus on the singers’ inner world rather than the external emotional connections.





file:///C:/Users/lynne/Documents/GitHub/project1/Output/project1.html

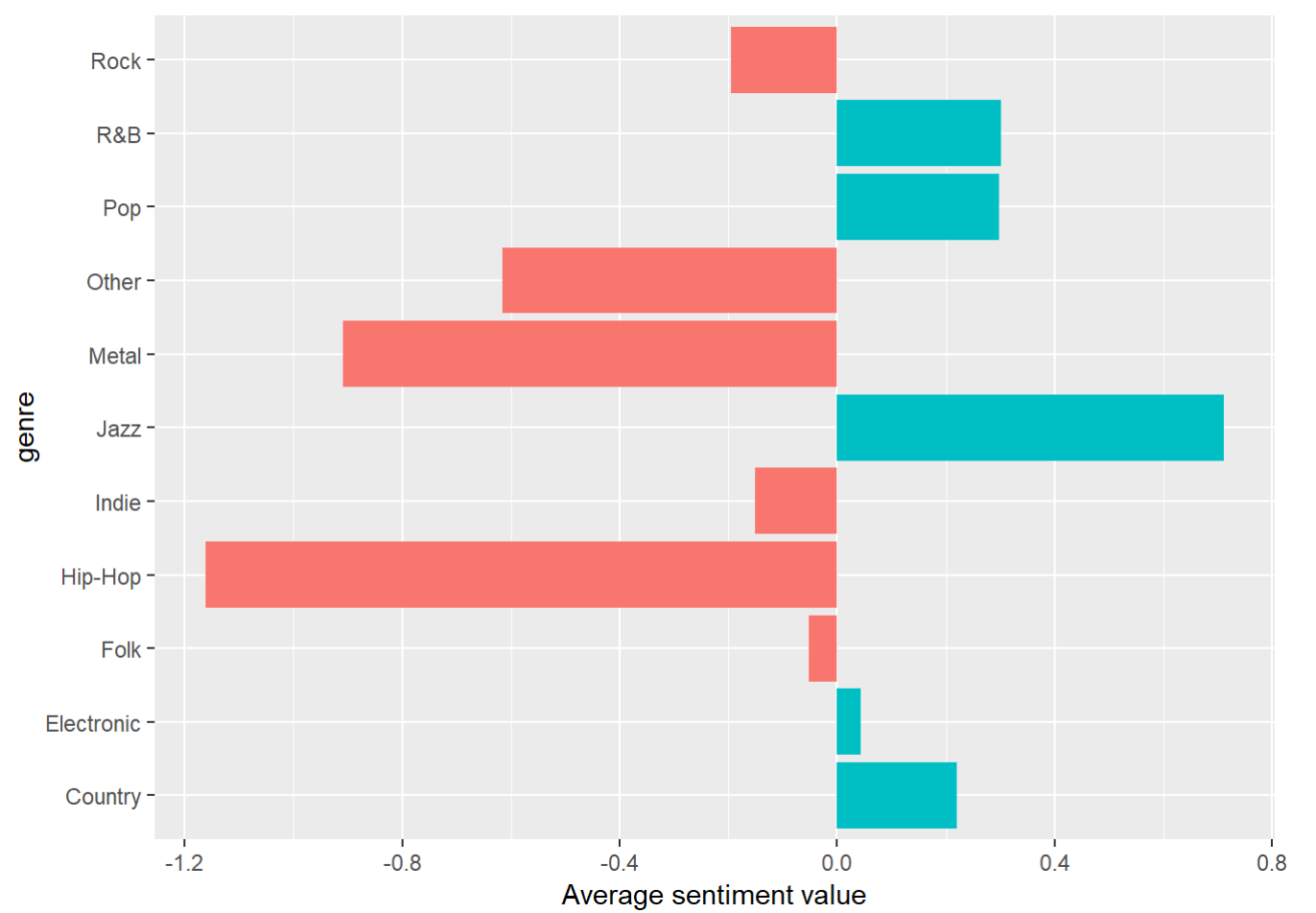
Sentiment Analysis

I would also like to examine which genre is the most positive or negative overall?

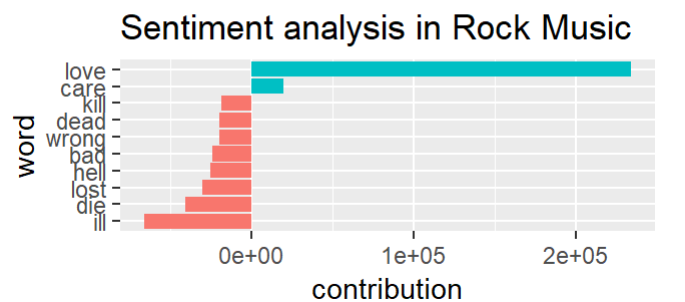
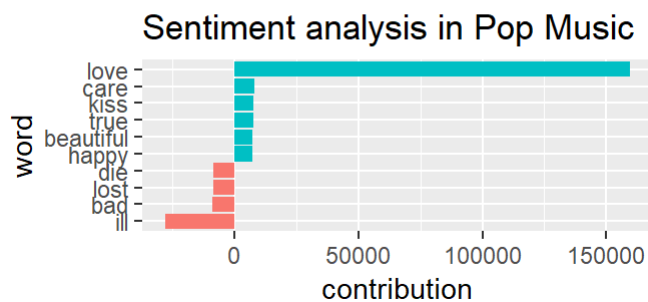
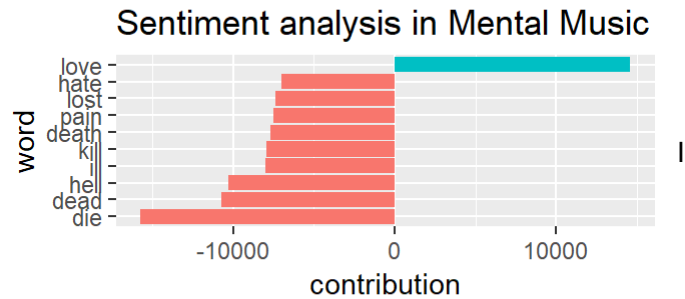
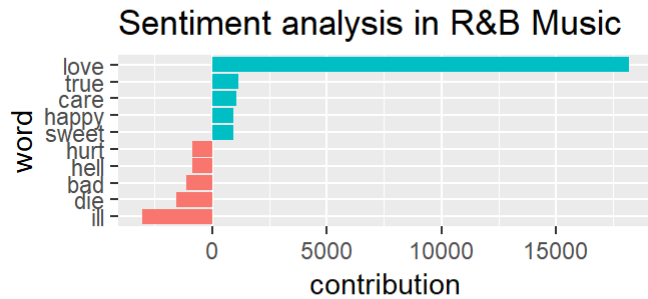
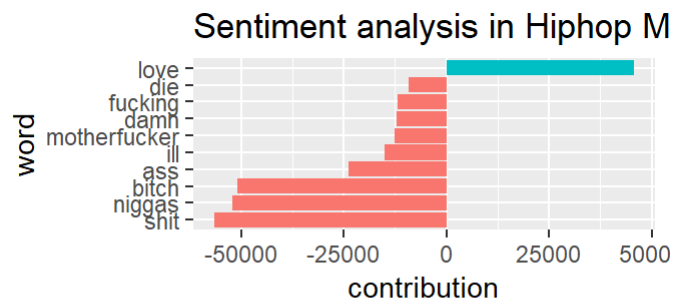
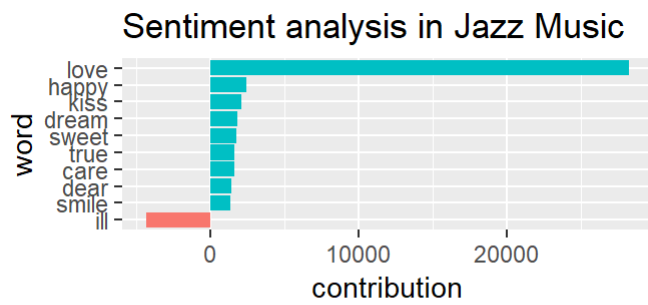
word	n
<chr>	<int>
love	187198
time	110131
youre	98054
baby	75341
ill	69885
day	63758
life	61115
heart	56928
ive	55756
night	53332

1-10 of 10,000 rows

Previous123456...1000Next



From the plot above, we can see the lyrics in Jazz music deliver the most positive values, and Hip-Hop music contains the most negative messages. These are corresponding to our findings in the word choice to express love.



take 6 different music genres, which are strongly affected by emotion, as examples, and look deeper to analyze which words contributed most to the positive and negative sentiment values. it turns out to be , jazz, R&B, and Pop music on the left contains frequent usage of “love”, “happy”, “true”, “care”, which are quite reasonable to be positive drivers. however, Hip-Hop, Mental, and Rock music mentions a lot about “ill”, “shit”, “die”, “hell” as the the strongest contributors to be negative lyrics.

Moreover, artists, as brilliant storytellers, also have fixed lyrics style, therefore, I list top 10 artists with the most negative songs.

genre <chr>	artist <chr>	sentiment <dbl>	words <int>
Country	dolly-parton	0.46835916	10714
Pop	bee-gees	0.40435459	9002
Rock	elton-john	0.07834702	7671
Rock	bob-dylan	-0.28251507	8127
Hip-Hop	chris-brown	-0.39434599	13937
Hip-Hop	chamillionaire	-0.74288039	9832
Hip-Hop	eminem	-1.56748237	25377
Hip-Hop	bone-thugs-n-harmony	-1.62518315	10920
Hip-Hop	busta-rhymes	-1.78863454	11227
Hip-Hop	dmx	-1.94753056	10063
1-10 of 10 rows			

Conclusions

1. Song lyrics changes with time, but “love” is the core emotion from past to present
2. Different genres of music have different patterns and word choice to express feelings.
3. Jazz music delivers the most positive messages and Hip-hop music delivers the most negative.