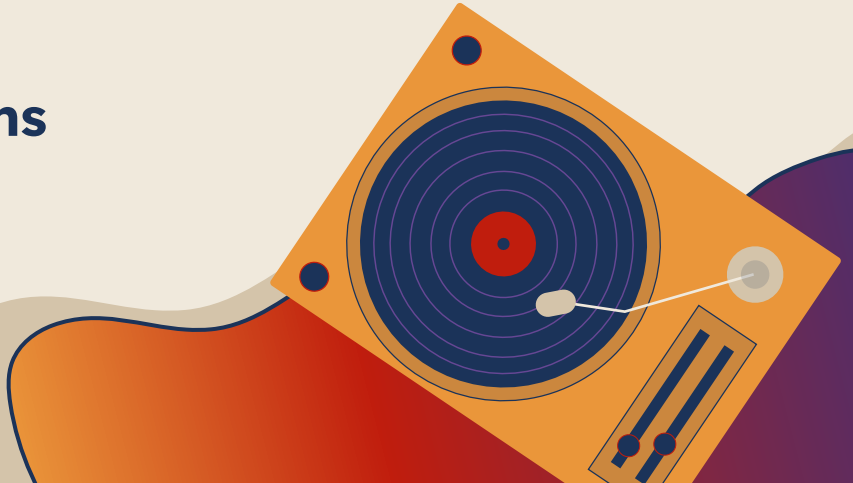
The Evolution of Trending Music: An Analysis

Abhinav Khanna (3036180545)
COMP2501 Project Presentation

Table of contents



01
Introduction

[redacted] 23/04/2024 23:31

[redacted]: Lonely No More (4) https://open.spotify.com/track/4faFAhOflLzhfJECveRwva?si=920feadec80244d2 vs
[redacted]: Out of My Mood (2) https://open.spotify.com/track/63qSlpq5oOEl2OdAopqpRh?si=8d5c9ae491eb447d

Lonely No More
Rob Thomas
PREVIEW

Out of My Mood
gate 문, brian mantra
PREVIEW E

2

[redacted]: Coming For You (1) https://open.spotify.com/track/2JDgJE2HdqghDvGysWbbjo?si=86e827f4c7994310
[redacted]: Laplace's Angel (3) https://open.spotify.com/track/1oHbgg5WYa7WOJC8VoBVYi?si=ad6a41fce8a04910

Coming For You
The Offspring
PREVIEW E

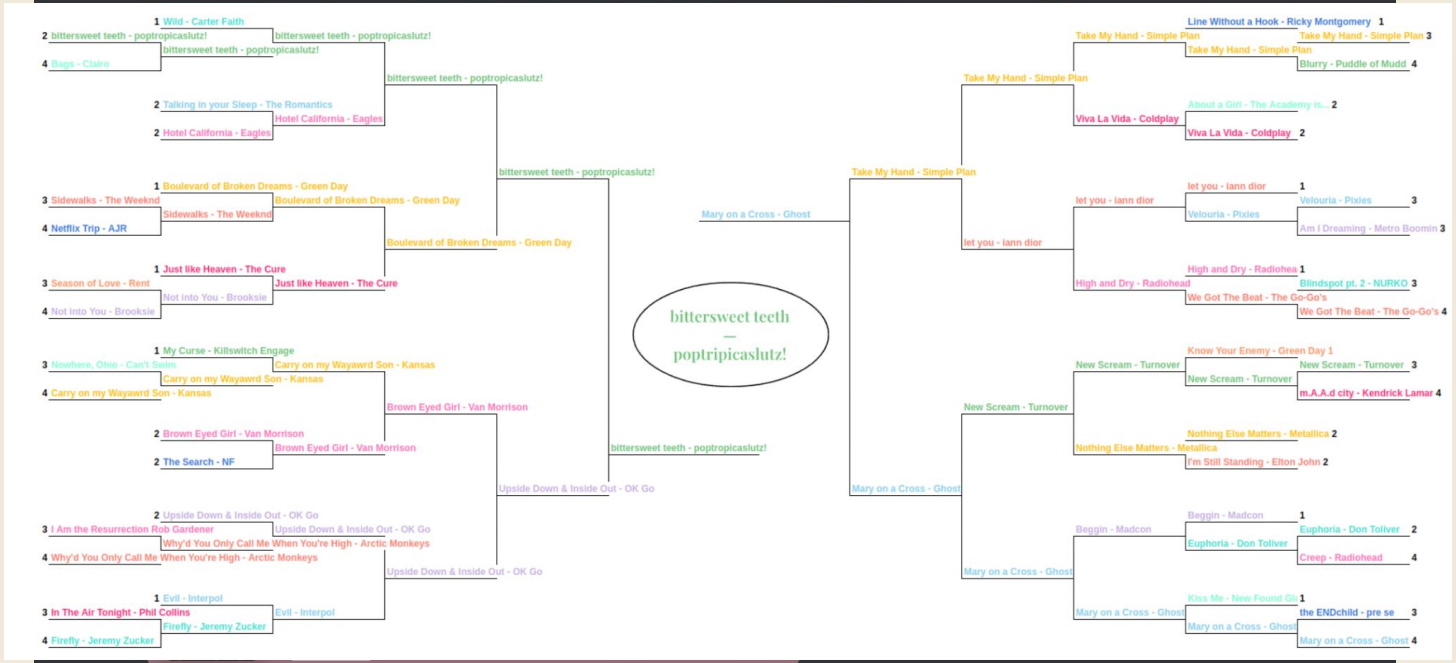
Laplace's Angel (Hurt People? Hurt People
Will Wood
PREVIEW

1 2

[redacted] 23/04/2024 23:31

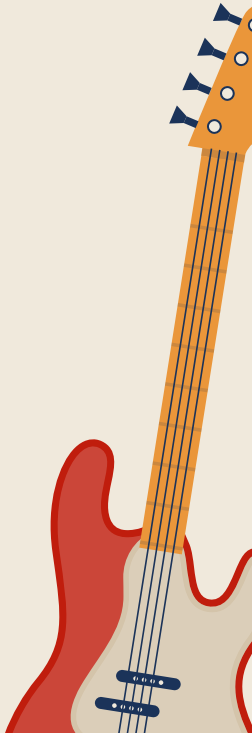
[redacted]: Lonely No More (4) https://open.spotify.com/track/4faFAhOflLzhfJECveRwva?si=920feadec80244d2 vs

[redacted]: Out of My Mood (2) https://open.spotify.com/track/63qSlpq5oOEl2OdAopqpRh?si=8d5c9ae491eb447d



1 2

Only listen to 2010’s music...



Only listen to 2010's music...

02

Research Questions

1) What features of trending music have changed across the last several decades?

1) What features of trending music have changed across the last several decades?

2) How do these changes in music reflect changes in our societies?

The Dataset

kaggle

Top 10000 Songs on Spotify 1960-Now

The best and biggest songs from ARIA & Billboard charts spanning 7 decades.

Spotify

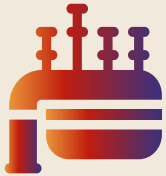
	Track.Name	Artist.Name.s.
1	Justified & Ancient - Stand by the Jams	The KLF
2	I Know You Want Me (Calle Ocho)	Pitbull
3	From the Bottom of My Broken Heart	Britney Spears
4	Apeman - 2014 Remastered Version	The Kinks
5	You Can't Always Get What You Want	The Rolling Stones
6	Don't Stop - 2004 Remaster	Fleetwood Mac
7	Eastside (with Halsey & Khalid)	benny blanco, Halsey, Khalid
8	Something About The Way You Look Tonight - Edit Version	Elton John
9	Juke Box Hero	Foreigner
10	Mercy	Shawn Mendes
11	It's Like That	Run-D.M.C., Jason Nevins
12	Here Without You	3 Doors Down
13	Listen to the Band - Single Version	The Monkees
14	With A Little Luck - Remastered 1993	Wings
15	Sing	Ed Sheeran
16	Mississippi	Pussycat
17	Flava	Nathaniel
18	Baby Sittin' Boogie - Radio Version	Buzz Clifford

	Album.Artist.Name.s.	Release.Year	Disc.Number	Track.Number	T
1	The KLF	1992	1	3	
2	Pitbull	2009	1	3	
3	Britney Spears	1999	1	6	
4	The Kinks	2014	1	11	
5	The Rolling Stones	1969	1	9	
6	Fleetwood Mac	1977	1	4	
7	benny blanco, Halsey, Khalid	2018	1	1	
8	Elton John	1997	1	1	
9	Foreigner	1981	1	2	
10	Shawn Mendes	2016	1	2	
11	Run-D.M.C.	2002	1	1	
12	3 Doors Down	2002	1	6	
13	The Monkees	2003	1	25	
14	Wings	1978	1	9	
15	Ed Sheeran	2014	1	3	
16	Pussycat	1976	1	8	
17	Nathaniel	2016	1	2	
18	Various Artists	2016	1	24	

Analysis is divided into:

Attribute Analysis

Based on various **song characteristic ratings** each decade



Genre Analysis

Based on **Spotify genre keywords** attached to songs each decade

03
Attribute Analysis

First, data cleansing...

```r
data <- read.csv("/Users/khanna1/Downloads/top_10000_1960-now.csv")

#class(data)

#colnames(data)

data_rlv <- data[, !(names(data) %in% c('Track.URI', 'Artist.URI.s.',
                                        'Album.Artist.URI.s.','Album.URI',
                                        'Album.Image.URL','Track.Preview.URL',
                                        'ISRC','Added.By','Added.At',
                                        'Copyrights'))]
```

```r
data_rlv$Album.Release.Date <- as.integer(substr(data_rlv$Album.Release.Date, 1, 4))


names(data_rlv)[names(data_rlv) == "Album.Release.Date"] <- "Release.Year"

data_rlv$Release.Year <- as.integer(data_rlv$Release.Year)

data_rlv <- data_rlv[complete.cases(data_rlv$Release.Year), ]
data_rlv <- data_rlv[data_rlv$Release.Year != 0, ]

#which(is.na(data_rlv$Release.Year))
#sum(is.na(data_rlv$Release.Year))

#data_rlv$Release.Year[1:20]

sorted_data <- data_rlv[order(data_rlv$Release.Year), ]
```

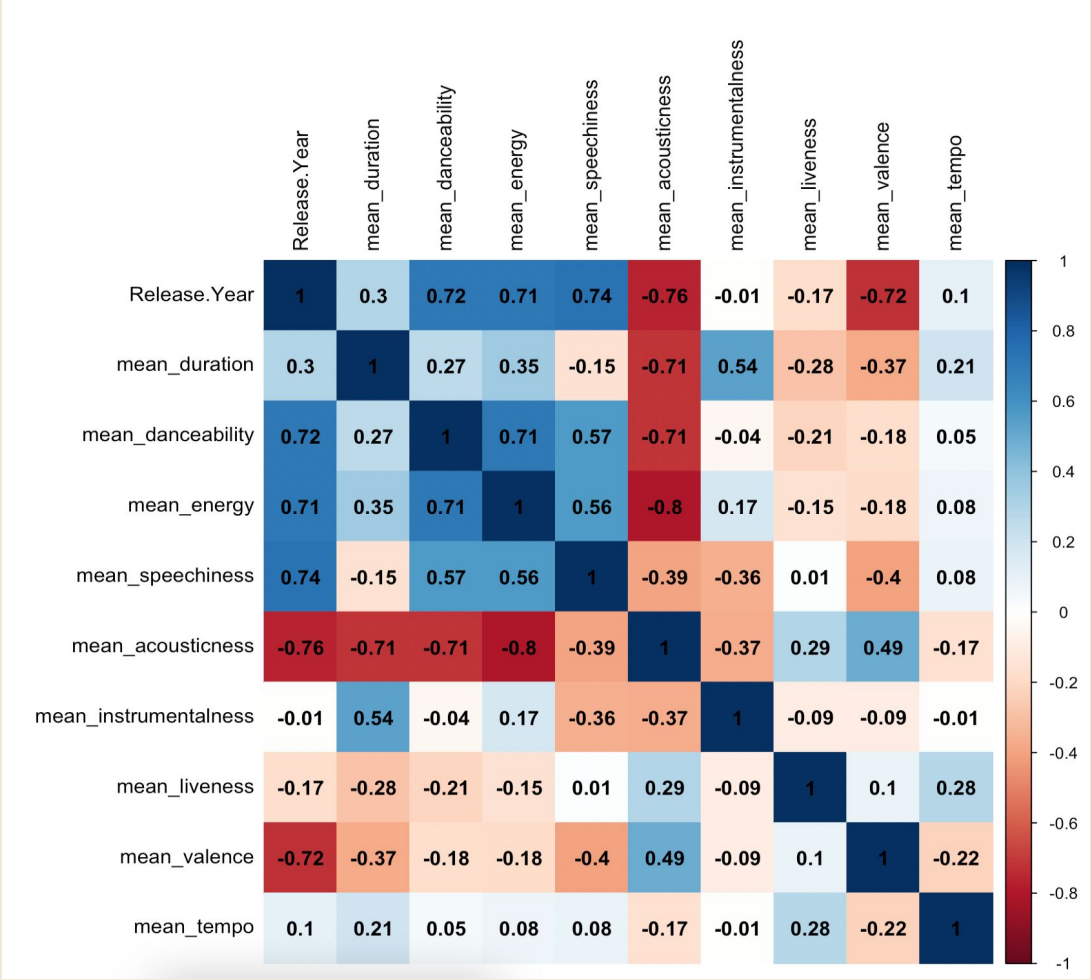
Album.Release.Date
1992-08-03
2009-10-23
1999-01-12
2014-10-20
1969-12-05
1977-02-04
2018-07-12
1997-01-01
1981
2016-09-23
2002-09-10
2002-11-12
2003
1978-03-31

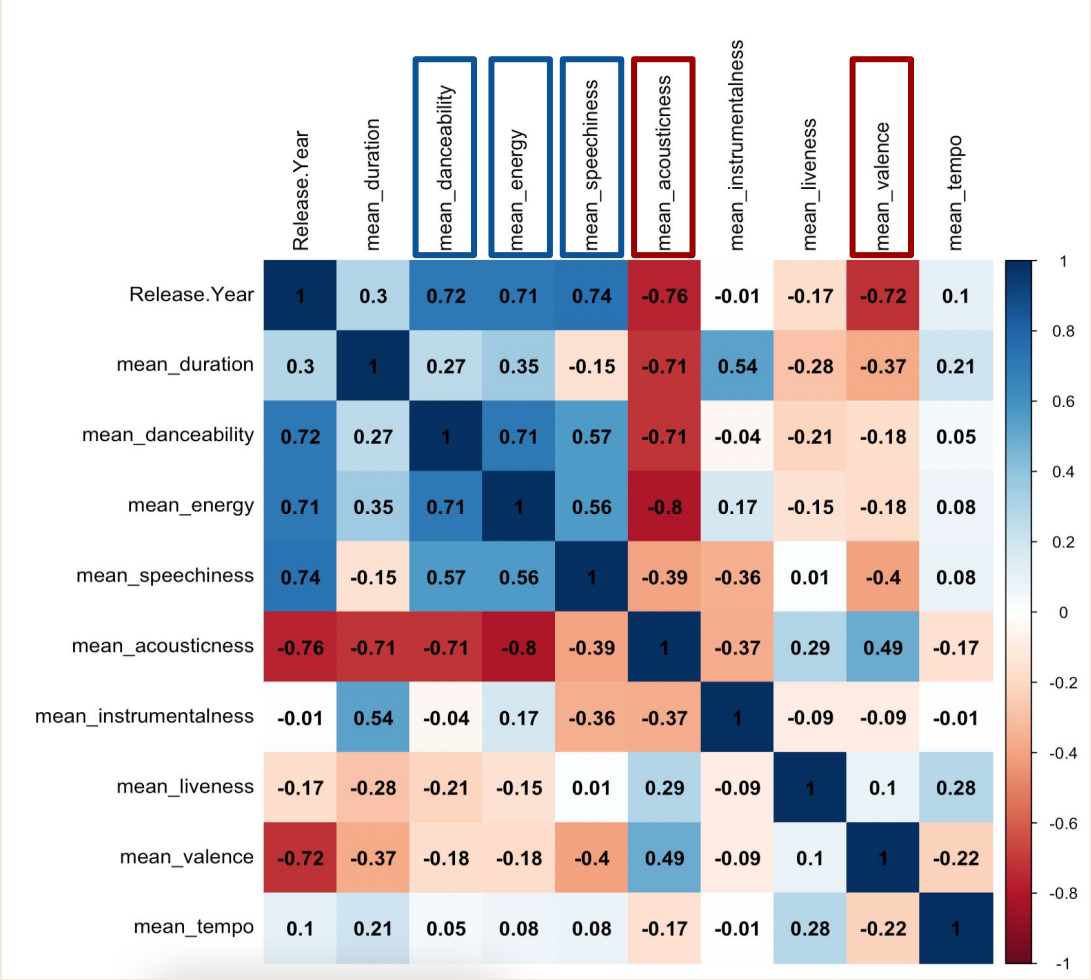
Group by Release Year, Visualise using Correlation Matrix

```r
data_grouped <- sorted_data |>
  group_by(Release.Year) |>
  summarise(mean_duration = mean(Track.Duration..ms.)/(1000*60),
            mean_danceability = mean(Danceability),
            mean_energy = mean(Energy),
            mean_speechiness = mean(Speechiness),
            mean_acousticness = mean(Acousticness),
            mean_instrumentalness = mean(Instrumentalness),
            mean_liveness = mean(Liveness),
            mean_valence = mean(Valence),
            mean_tempo = mean(Tempo))
```

```r
cor_data <- cor(data_grouped)

corrplot(cor_data, method = "color", type = "full", tl.col = "black",
         addCoef.col = "black")
```


Release.Year
mean_duration
mean_danceability
mean_energy
mean_speechiness
mean_acousticness
mean_instrumentalness
mean_liveness
mean_valence
mean_tempo
Release.Year 1 0.3 0.72 0.71 0.74 -0.76 -0.01 -0.17 -0.72 0.1
mean_duration 0.3 1 0.27 0.35 -0.15 -0.71 0.54 -0.28 -0.37 0.21
mean_danceability 0.72 0.27 1 0.71 0.57 -0.71 -0.04 -0.21 -0.18 0.05
mean_energy 0.71 0.35 0.71 1 0.56 -0.8 0.17 -0.15 -0.18 0.08
mean_speechiness 0.74 -0.15 0.57 0.56 1 -0.39 -0.36 0.01 -0.4 0.08
mean_acousticness -0.76 -0.71 -0.71 -0.8 -0.39 1 -0.37 0.29 0.49 -0.17
mean_instrumentalness -0.01 0.54 -0.04 0.17 -0.36 -0.37 1 -0.09 -0.09 -0.01
mean_liveness -0.17 -0.28 -0.21 -0.15 0.01 0.29 -0.09 1 0.1 0.28
mean_valence -0.72 -0.37 -0.18 -0.18 -0.4 0.49 -0.09 0.1 1 -0.22
mean_tempo 0.1 0.21 0.05 0.08 0.08 -0.17 -0.01 0.28 -0.22 1
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1


Release.Year
mean_duration
mean_danceability
mean_energy
mean_speechiness
mean_acousticness
mean_instrumentalness
mean_liveness
mean_valence
mean_tempo
Release.Year 1 0.3 0.72 0.71 0.74 -0.76 -0.01 -0.17 -0.72 0.1
mean_duration 0.3 1 0.27 0.35 -0.15 -0.71 0.54 -0.28 -0.37 0.21
mean_danceability 0.72 0.27 1 0.71 0.57 -0.71 -0.04 -0.21 -0.18 0.05
mean_energy 0.71 0.35 0.71 1 0.56 -0.8 0.17 -0.15 -0.18 0.08
mean_speechiness 0.74 -0.15 0.57 0.56 1 -0.39 -0.36 0.01 -0.4 0.08
mean_acousticness -0.76 -0.71 -0.71 -0.8 -0.39 1 -0.37 0.29 0.49 -0.17
mean_instrumentalness -0.01 0.54 -0.04 0.17 -0.36 -0.37 1 -0.09 -0.09 -0.01
mean_liveness -0.17 -0.28 -0.21 -0.15 0.01 0.29 -0.09 1 0.1 0.28
mean_valence -0.72 -0.37 -0.18 -0.18 -0.4 0.49 -0.09 0.1 1 -0.22
mean_tempo 0.1 0.21 0.05 0.08 0.08 -0.17 -0.01 0.28 -0.22 1
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

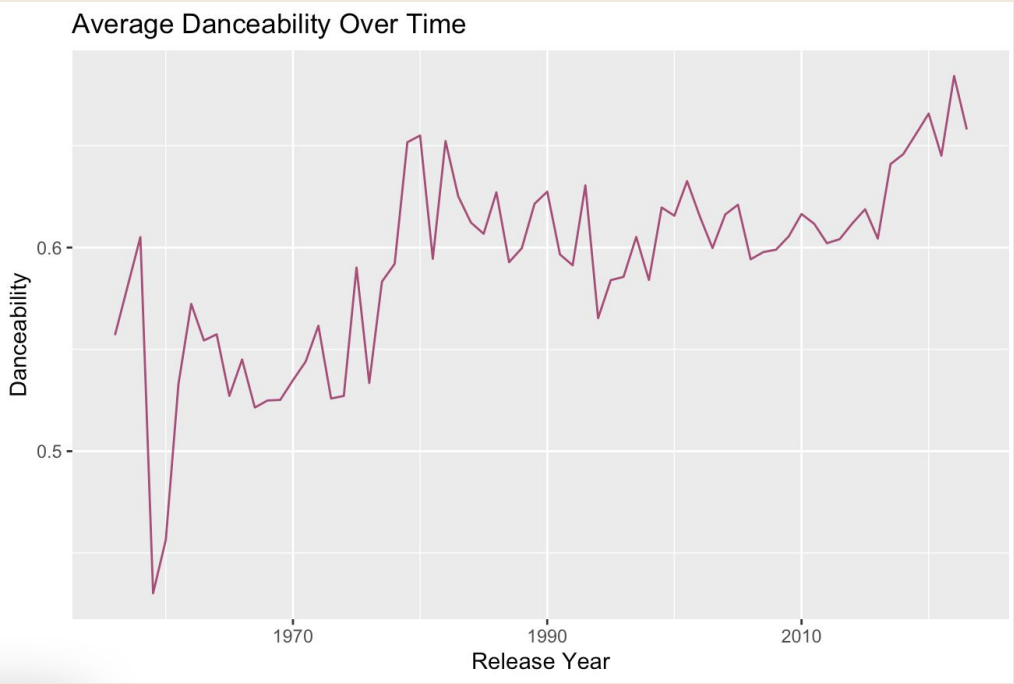
Consider sample size!

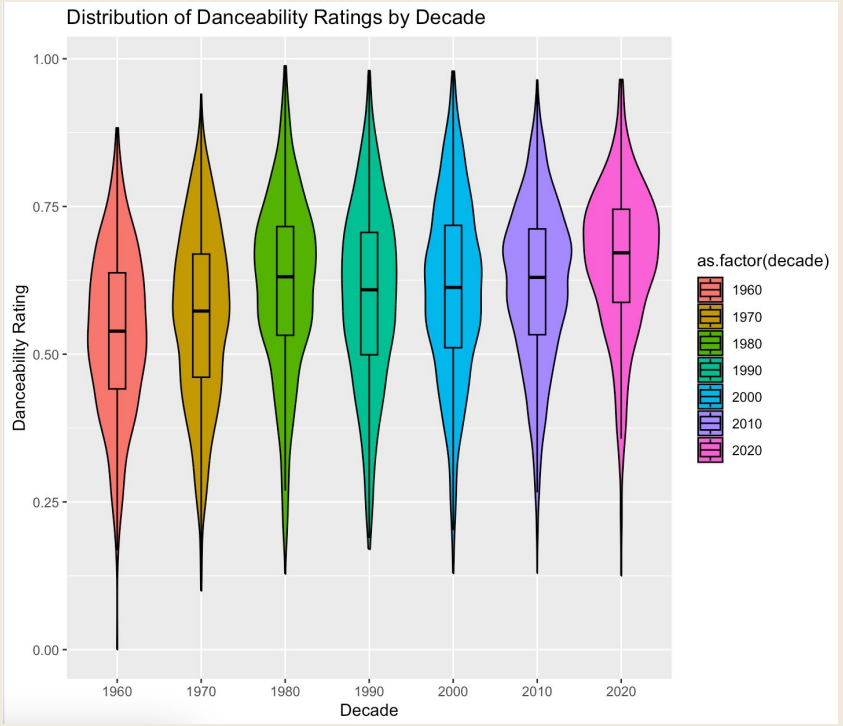
```
song_counts_decade = table(sorted_data$decade)
song_counts_decade
```

```
1950 1960 1970 1980 1990 2000 2010 2020
  16  442  747  931 1303 2551 3450  556
```

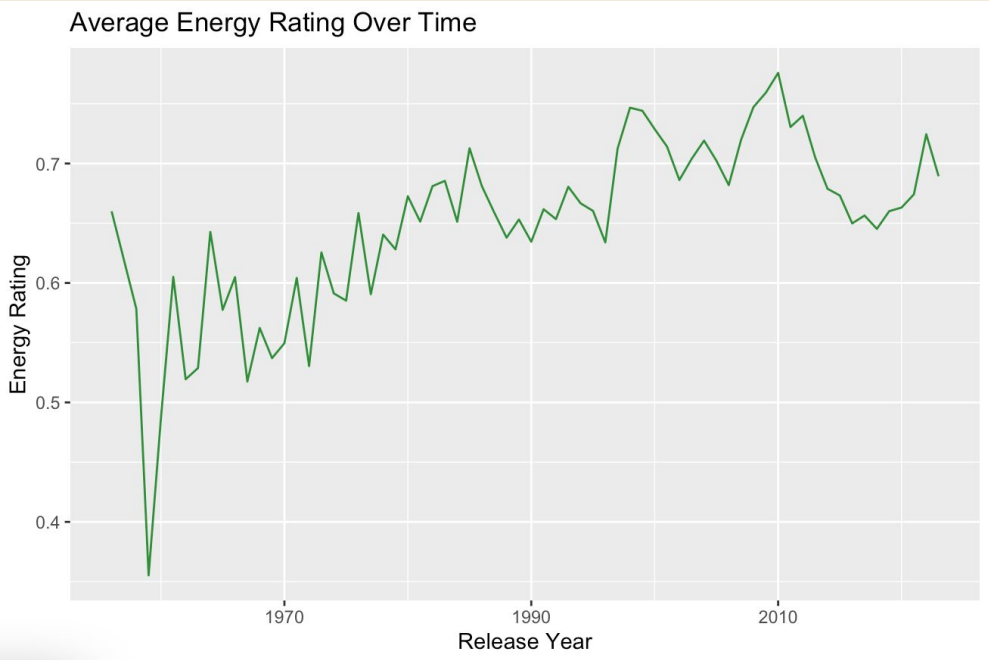
```
filtered_songs <- sorted_data[sorted_data$decade != 1950, ]
```

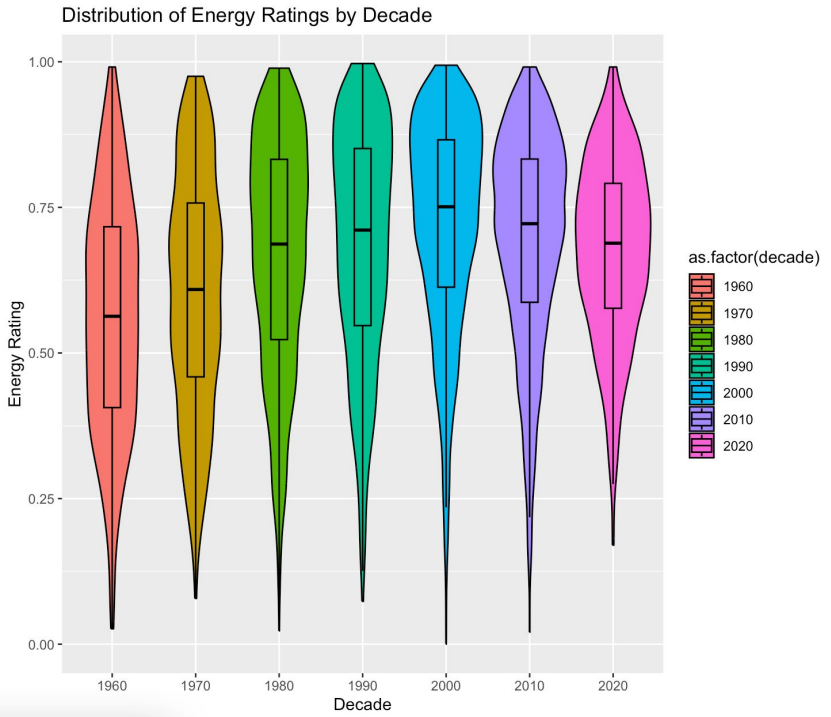
Danceability



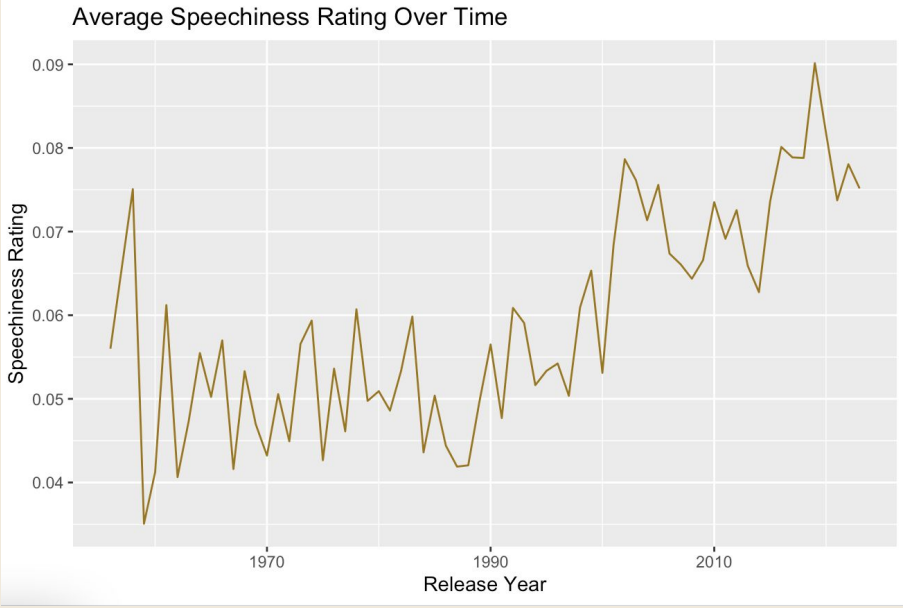


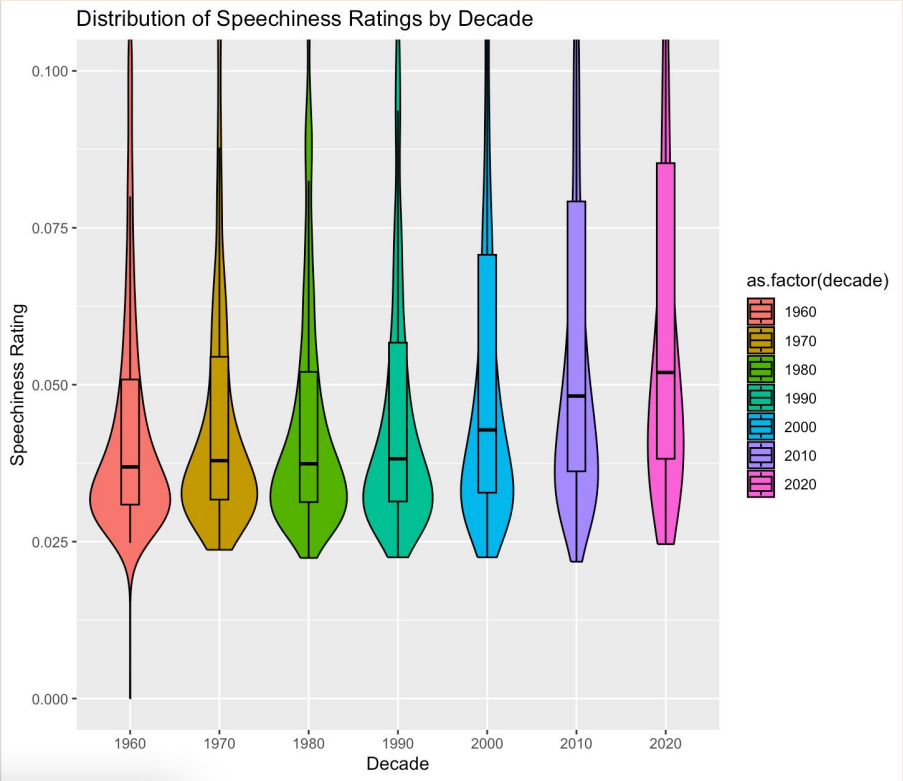
Energy


Average Energy Rating Over Time
Energy Rating
0.7
0.6
0.5
0.4
1970
1990
2010
Release Year

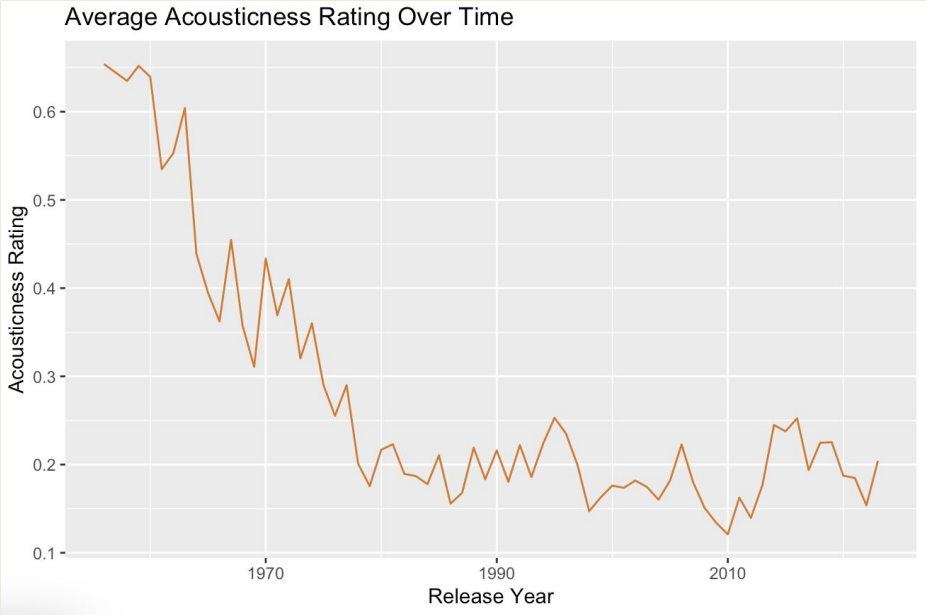

Distribution of Energy Ratings by Decade
Energy Rating
1.00
0.75
0.50
0.25
0.00
1960
1970
1980
1990
2000
2010
2020
Decade
as.factor(decade)
1960
1970
1980
1990
2000
2010
2020

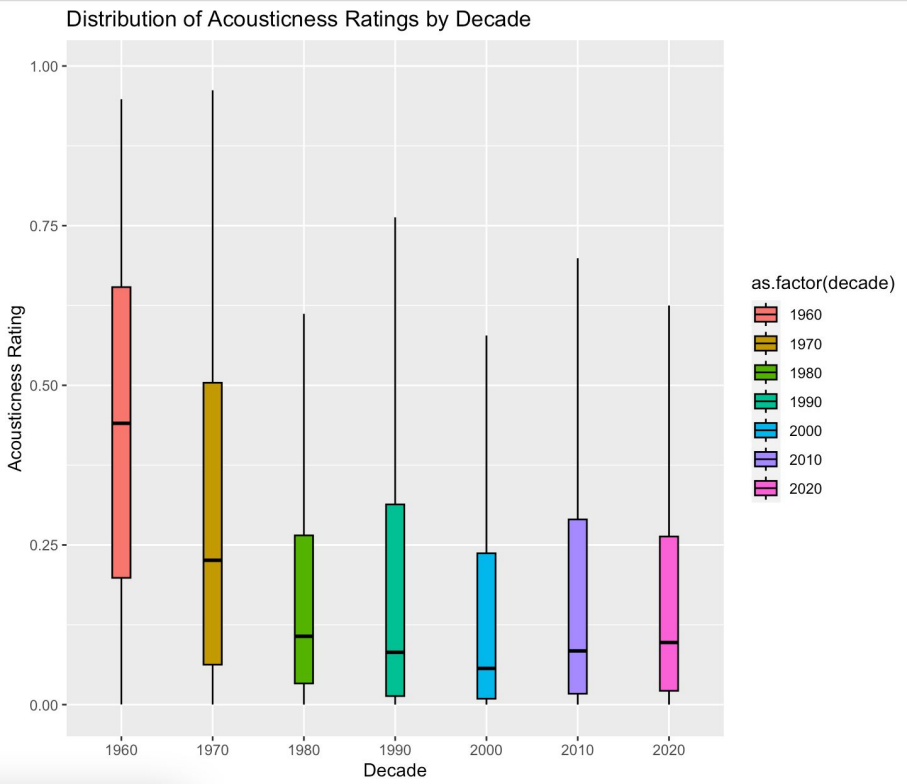
Speechiness



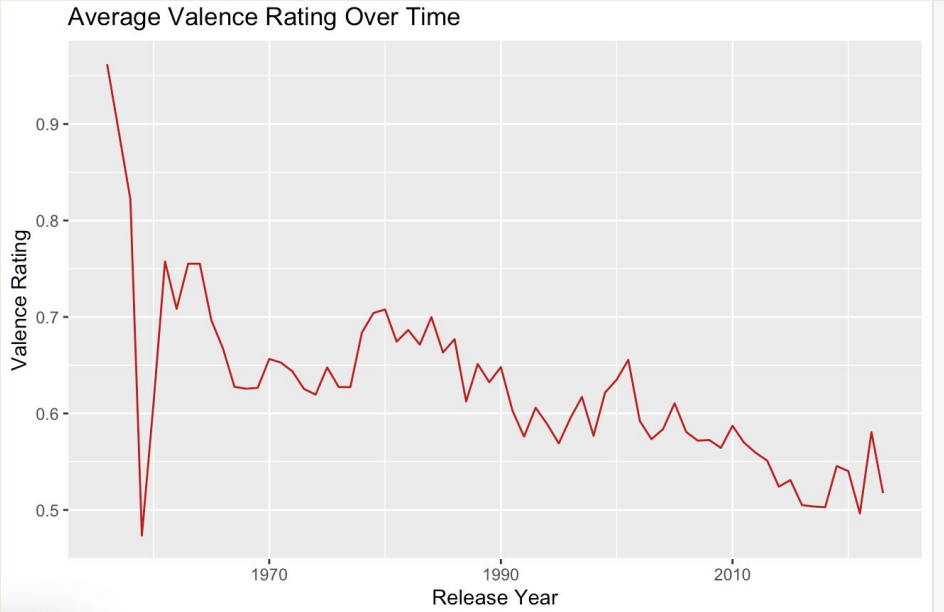


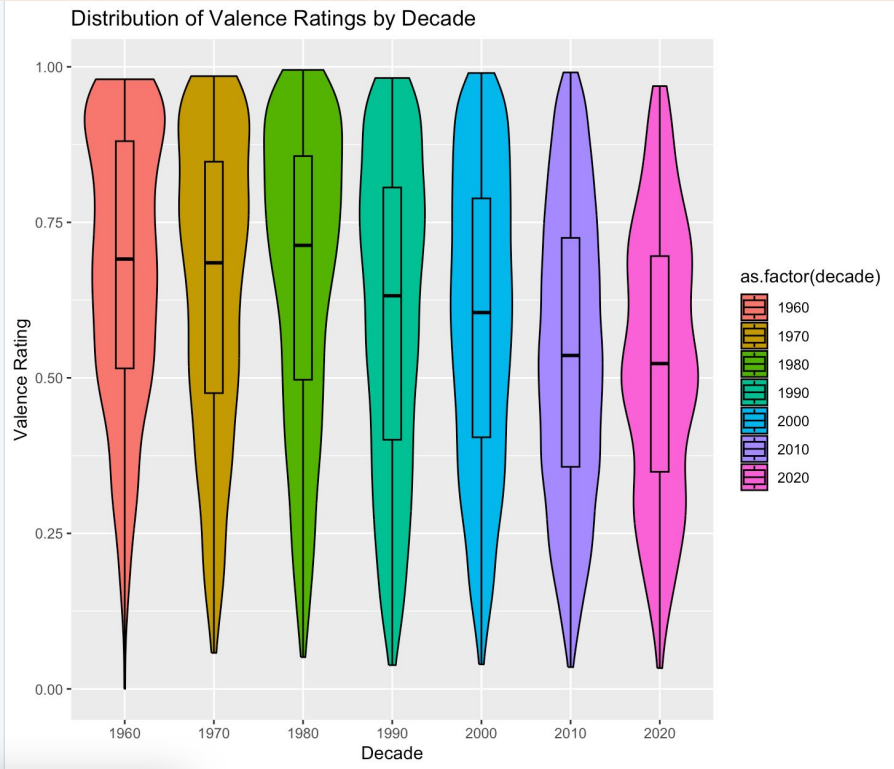
Acousticness





Valence


Average Valence Rating Over Time
Valence Rating
0.9
0.8
0.7
0.6
0.5
1970
1990
2010
Release Year


Distribution of Valence Ratings by Decade
Valence Rating
1.00
0.75
0.50
0.25
0.00
1960
1970
1980
1990
2000
2010
2020
Decade
as.factor(decade)
1960
1970
1980
1990
2000
2010
2020

04

Genre Analysis

Using Artist Genres Column:

```
> sorted_data$Artist.Genres[1:20]
 [1] "rock-and-roll,rockabilly"
 [2] "doo-wop,rhythm and blues"
 [3] "adult standards,easy listening,lounge,rock-and-roll,vocal jazz"
 [4] "classic rock,folk rock,mellow gold,rock-and-roll,rockabilly,singer-songwriter,doo-wop,rock-and-roll,rockabilly"
 [5] "doo-wop,rock-and-roll"
 [6] "classic rock,folk rock,mellow gold,rock-and-roll,rockabilly,singer-songwriter,doo-wop,rock-and-roll,rockabilly"
 [7] "rock-and-roll,rockabilly"
 [8] "classic rock,folk rock,mellow gold,rock-and-roll,rockabilly,singer-songwriter,doo-wop,rock-and-roll,rockabilly"
 [9] "chicago soul,classic soul,rock-and-roll,soul"
[10] "rock-and-roll,rockabilly"
[11] "rock-and-roll,rockabilly"
[12] "rock-and-roll,rockabilly"
[13] ""
[14] "adult standards,easy listening,lounge,rock-and-roll,vocal jazz"
[15] "adult standards,brill building pop,easy listening,rock-and-roll,rockabilly"
[16] "bebop,jazz,jazz quartet"
[17] "chicago soul,classic soul,rock-and-roll,soul"
[18] "adult standards,folk rock,mellow gold,rock-and-roll,rockabilly,sunshine pop"
```

Separate by comma, count the frequencies

```
song_genre_df <- filtered_songs |>
  separate_rows(Artist.Genres, sep = ",\\s*") |>
  mutate(Artist.Genres = trimws(Artist.Genres))
```

```
word_freq <- song_genre_df |>
  group_by(decade) |>
  count(Artist.Genres, sort = TRUE)

top_words <- word_freq |>
  group_by(decade) |>
  arrange(decade, desc(n)) |>
  top_n(10)
```

```
> top_words
# A tibble: 78 × 3
# Groups:   decade [7]
   decade Artist.Genres          n
    <dbl> <chr>               <int>
 1   1960 classic rock          152
 2   1960 rock                  113
 3   1960 folk rock             112
 4   1960 british invasion      106
 5   1960 rock-and-roll          97
 6   1960 psychedelic rock       93
```

Genre Frequencies

1960's

TOTAL:	2104
"classic rock"	152
"rock"	113
"folk rock"	112
"british invasion"	106
"rock-and-roll"	97
"psychedelic rock"	93
"mellow gold"	87
"soul"	73
"adult standards"	72
"singer-songwriter"	70

1970's

TOTAL:	3240
"classic rock"	281
"soft rock"	280
"mellow gold"	263
"rock"	247
"album rock"	206
"singer-songwriter"	127
"hard rock"	113
"folk rock"	111
"glam rock"	99
"heartland rock"	92

1980's

TOTAL:	3886
"soft rock"	287
"new wave pop"	280
"rock"	271
"new romantic"	216
"classic rock"	198
"album rock"	195
"mellow gold"	166
"new wave"	154
"australian rock"	143
"synthpop"	128

1990's

TOTAL:	4703
"rock"	222
"new wave pop"	149
"dance pop"	139
"permanent rock"	128
"australian rock"	123
"soft rock"	115
"pop rock"	110
"alternative rock"	96
"post-grunge"	93
"mellow gold"	90

Genre Frequencies

1960's		1970's		1980's		1990's	
TOTAL:	**2104**	**TOTAL:**	**3240**	**TOTAL:**	**3886**	**TOTAL:**	**4703**
"classic rock"	152	"classic rock"	281	"soft rock"	287	"rock"	222
"rock"	113	"soft rock"	280	"new wave pop"	280	"new wave pop"	149
"folk rock"	112	"mellow gold"	263	"rock"	271	"dance pop"	139
"british invasion"	[illegible]	"rock"	[illegible]	"new romantic"	[illegible]	"permanent rock"	[illegible]
"rock-and-roll"	[illegible]	"album rock"	[illegible]	"classic rock"	[illegible]	"australian rock"	[illegible]
"psychedelic rock"	[illegible]	"singer-songwriter"	[illegible]	"album rock"	[illegible]	"soft rock"	[illegible]
"mellow gold"	87	"hard rock"	[illegible]	"mellow gold"	[illegible]	"pop rock"	[illegible]
"soul"	73	"folk rock"	111	"new wave"	154	"alternative rock"	96
"adult standards"	72	"glam rock"	99	"australian rock"	143	"post-grunge"	93
"singer-songwriter"	70	"heartland rock"	92	"synthpop"	128	"mellow gold"	90

Genre Frequencies

1960's

TOTAL:	100%
"classic rock"	7.22%
"rock"	5.37%
"folk rock"	5.32%
"british invasion"	5.04%
"rock-and-roll"	4.61%
"psychedelic rock"	4.42%
"mellow gold"	4.13%
"soul"	3.47%
"adult standards"	3.42%
"singer-songwriter"	3.33%

1970's

TOTAL:	100%
"classic rock"	8.67%
"soft rock"	8.64%
"mellow gold"	8.12%
"rock"	7.62%
"album rock"	6.36%
"singer-songwriter"	3.92%
"hard rock"	3.49%
"folk rock"	3.43%
"glam rock"	3.06%
"heartland rock"	2.84%

1980's

TOTAL:	100%
"soft rock"	7.39%
"new wave pop"	7.21%
"rock"	6.97%
"new romantic"	5.56%
"classic rock"	5.10%
"album rock"	5.02%
"mellow gold"	4.27%
"new wave"	3.96%
"australian rock"	3.68%
"synthpop"	3.29%

1990's

TOTAL:	100%
"rock"	4.72%
"new wave pop"	3.17%
"dance pop"	2.96%
"permanent rock"	2.72%
"australian rock"	2.62%
"soft rock"	2.45%
"pop rock"	2.34%
"alternative rock"	2.04%
"post-grunge"	1.98%
"mellow gold"	1.91%

Genre Frequencies

2000's

TOTAL:	100%
"dance pop"	6.20%
"pop"	5.49%
"rock"	3.22%
"urban contemporary"	2.43%
"australian rock"	2.24%
"r&b"	2.08%
"australian pop"	2.07%
"pop rap"	2.06%
"pop rock"	2.04%
"rap"	1.88%

2010's

TOTAL:	100%
"pop"	13.88%
"dance pop"	5.63%
"pop dance"	2.74%
"edm"	2.65%
"uk pop"	2.44%
"rap"	2.42%
"pop rap"	2.27%
"australian pop"	2.19%
"hip hop"	1.48%
"modern rock"	1.42%

2020's

TOTAL:	100%
"pop"	18.65%
"dance pop"	5.83%
"uk pop"	4.68%
"pop dance"	3.94%
"rap"	3.42%
"edm"	2.79%
"uk dance"	2.47%
"australian pop"	2.31%
"canadian pop"	1.79%
"alt z"	1.68%

05
Conclusions

Music has evolved to be:

- More **energetic** and **danceable**
- Greater presence of **spoken words**
- **Less positive** sounding...

What does this say about society?

- **Pop music** is the new **king**
- We appreciate **messages** in modern music
- Greater preference towards music that evokes **negative emotions**, but we still love a **bop**

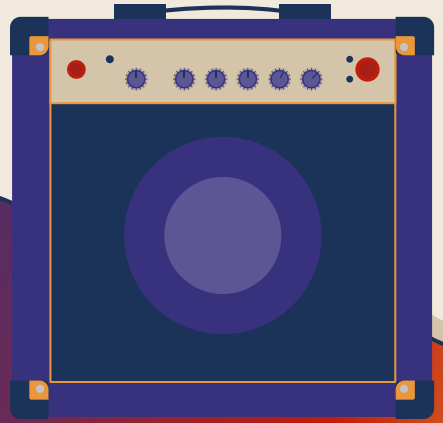
Project Limitations

1) The **most trending** music may not be the best **critically-acclaimed**
2) Analysis does not distinguish between **relative popularity** of tracks
3) Dataset is only taken from **Spotify**, only includes tracks available on the platform

Possible Extensions

1) Consider relative popularities **per country** along with **Happiness Index**
2) Use **ML** to predict time periods based on attributes or genre

Thank you!



CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**