Exploring the Sentiment of Lyrics from Billboard Top 100 1960-2015

Tuomo Kareoja

December 17, 2019

Agenda

WebScraping
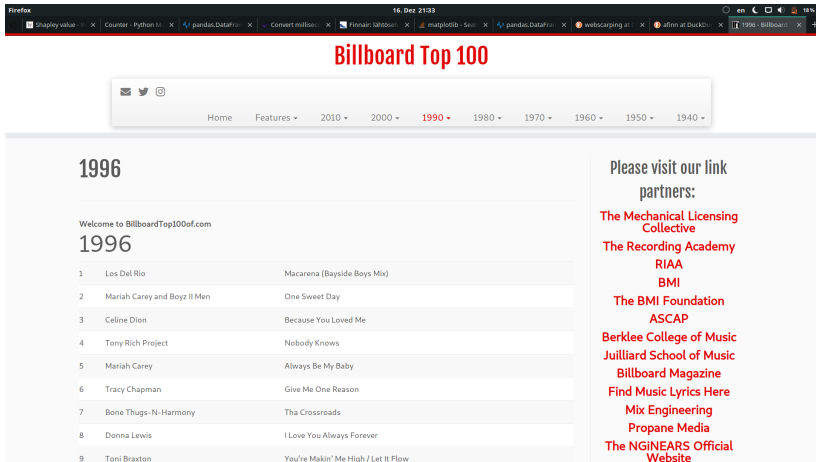
▶ Reading the contents from HTML tables with the BeautifulSoup package

Lyrics From APIs

▶ Used three different APIs with the Requests package
1. Mourits Lyrics (best, but so much down!)
2. Genius Lyrics
3. ChartLyrics

Lyrics From APIs

- Used three different APIs with the Requests package
  1. Mourits Lyrics (best, but so much down!)
  2. Genius Lyrics
  3. ChartLyrics
- What you get differs from site to site as the lyrics are added by users
  - If the lyric is weird, try a different API (no long bible quotes!)

Get Track Details From Spotify

- ▶ Spotify has a really nice API for getting information about the actual audio (e.g. length, danceability, tempo, loudness...)

Get Track Details From Spotify

- Spotify has a really nice API for getting information about the actual audio (e.g. length, danceability, tempo, loudness...)
- You need to register an app, and the API, is bit fiddly...

Get Track Details From Spotify

- ▶ Spotify has a really nice API for getting information about the actual audio (e.g. length, danceability, tempo, loudness...)
- ▶ You need to register an app, and the API, is bit fiddly...
- ▶ But of course there is a python package does most of the work for you! (spotipy)

Get Track Details From Spotify

- ▶ Spotify has a really nice API for getting information about the actual audio (e.g. length, danceability, tempo, loudness...)
- ▶ You need to register an app, and the API, is bit fiddly...
- ▶ But of course there is a python package does most of the work for you! (spotipy)
- ▶ Still need to first find the song ID and then find the audio features with this, so multiple calls needed

Cleaning the Lyrics

▶ There are bots and trolling

Rihanna - You Da One
Hello! dafdcdk interesting dafdcdk site! I'm really like it! Very, very dafdcdk good!

Cleaning the Lyrics

- ▶ There are bots and trolling

Rihanna - You Da One
Hello! dafdcdk interesting dafdcdk site! I'm really like it! Very, very dafdcdk good!

- ▶ Non-standard spelling is commonplace

Dr. Dre - Nothing But a G Thang
...yo' attention Mobbin like a muh'fucker, but I ain't lynchin Droppin the funky shit that's makin the sucka niggaz mumble...

Cleaning the Lyrics

- There are bots and trolling

Rihanna - You Da One
Hello! dafdcdk interesting dafdcdk site! I'm really like it! Very, very
dafdcdk good!

- Non-standard spelling is commonplace

Dr. Dre - Nothing But a G Thang
...yo' attention Mobbin like a muh'fucker, but I ain't lynchin Droppin
the funky shit that's makin the sucka niggaz mumble...

- And so many other problems too...

Is It Instrumental?

- Try to find mention of song being instrumental from lyrics
  - "Instrumental, NO LYRICS"
  - But what if there is an instrumental part marked in the lyrics? ("Instrumental Interlude")

Is It Instrumental?

- Try to find mention of song being instrumental from lyrics
  - "Instrumental, NO LYRICS"
  - But what if there is an instrumental part marked in the lyrics? ("Instrumental Interlude")
- Using very short lyrics is better
  - "Ah ha ha ha ha ha ha ha ha ha ha ha, wipe out", Surfaris - Wipe Out, 1963

Is It Instrumental?

- Try to find mention of song being instrumental from lyrics
  - "Instrumental, NO LYRICS"
  - But what if there is an instrumental part marked in the lyrics? ("Instrumental Interlude")
- Using very short lyrics is better
  - "Ah ha ha ha ha ha ha ha ha ha ha ha, wipe out", Surfaris - Wipe Out, 1963
- Use Spotify's instrumentalness
  - Gives and estimated probability of the song being instrumental

Is It In English?

- Billboard contains also some Spanish, French and Italian(!) songs

Is It In English?

- ▶ Billboard contains also some Spanish, French and Italian(!) songs
- ▶ These can be recognized with the langdetect package

Is It In English?

- Billboard contains also some Spanish, French and Italian(!) songs
- These can be recognized with the langdetect package
- ...but some songs just have spanish parts (or trolling or mistakes)

Rod Stewart - Do You Think I'm Sexy
...Cariño, cariño Ella se sienta sola esperando propuestas Es está muy nervioso evitando todas las preguntas Los labios de él están secos...

Measuring Sentiment

Afinn
- Lexicon of English terms manually rated for valence with an integer between -5 (negative) and 5 (positive)
  - afinn.score('Terrible') = -3.0
  - afinn.score('Make Berlin great again!') = 3.0

Measuring Sentiment

Afinn

- Lexicon of English terms manually rated for valence with an integer between -5 (negative) and 5 (positive)
  - afinn.score('Terrible') = -3.0
  - afinn.score('Make Berlin great again!') = 3.0

Emolex

- Lexicon of English words manually rated for associations with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and two sentiments (negative and positive)
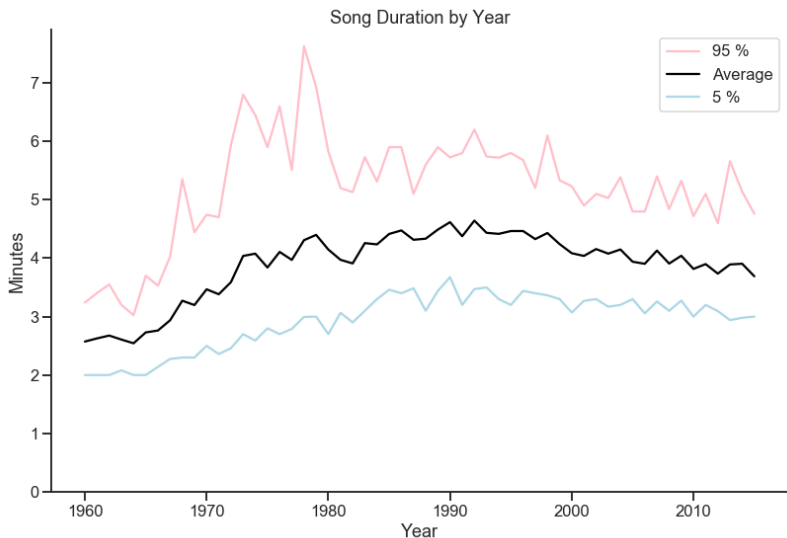
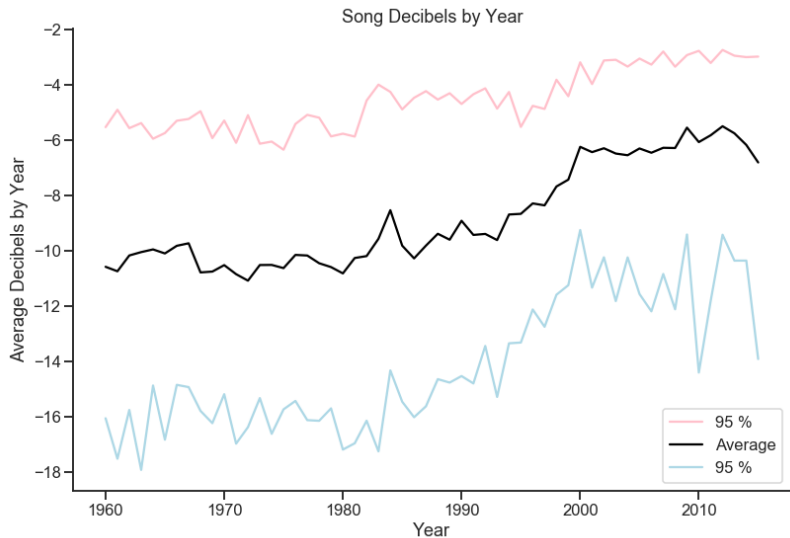| abnormal | disgust | 1 |
| abnormal | negative | 1 |

Results

Audio Features

Technology Has Pushed Song Duration Up, But Not Anymore

Song Duration by Year

Songs are Getting Louder

Song Decibels by Year

More Words per Minute (Rap's Influence?)

Average Words per Minute by Year
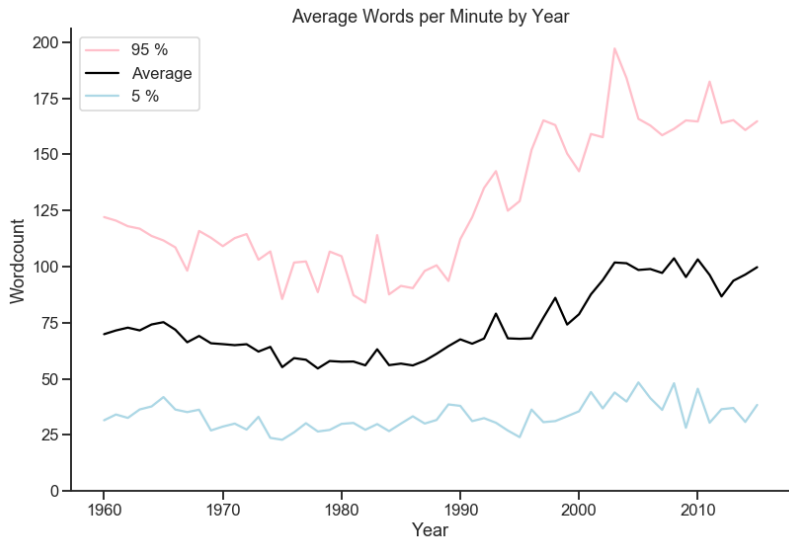
Death of Acousticness/Rise of Dance Music
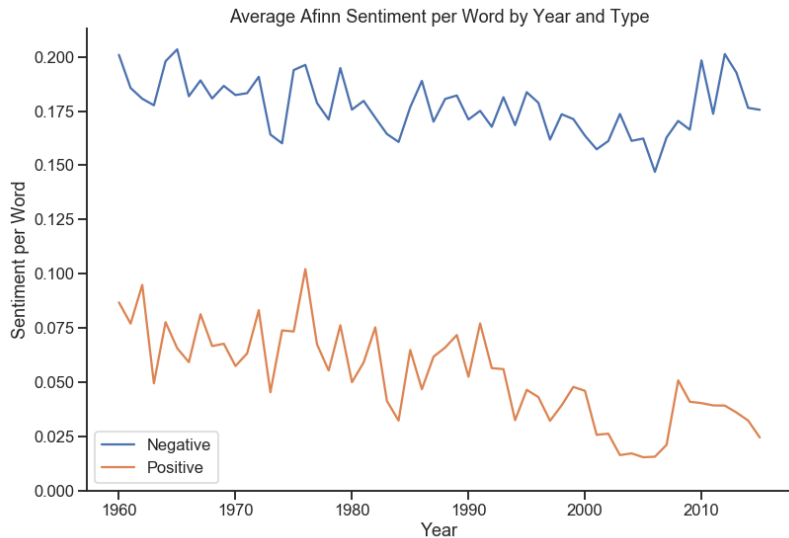

Average Value of Spotify Audio Features by Year
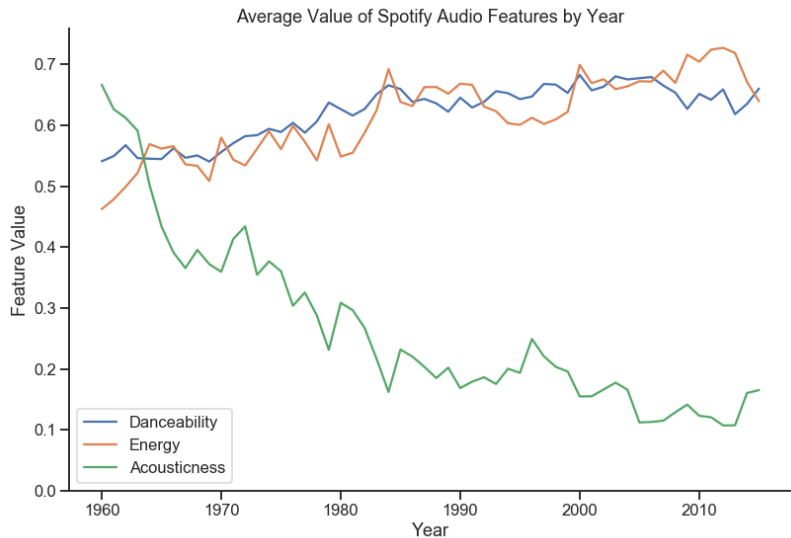
Audio Features Show Decreasing Positivity
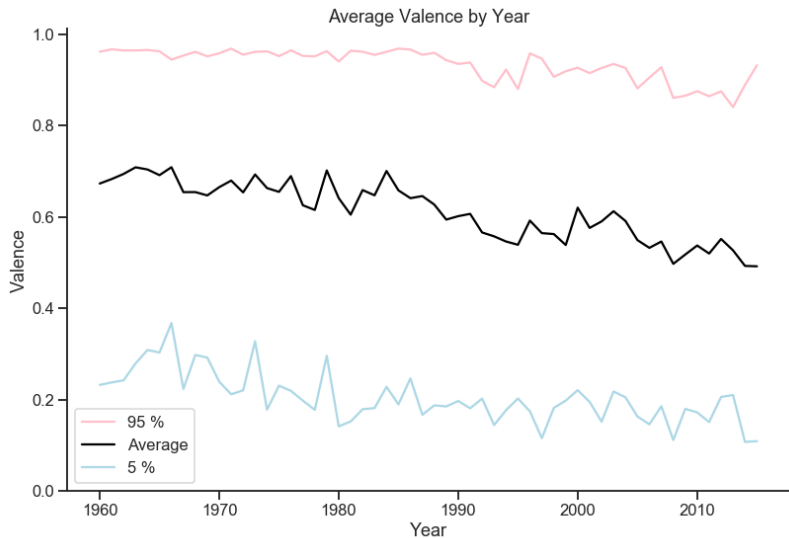
Lyrics' Sentiment

Afinn Positive Sentiment is in Decline

Average Afinn Sentiment per Word by Year and Type

NRC Shows the Same

Average Overall NRC Sentiment per Word by Year
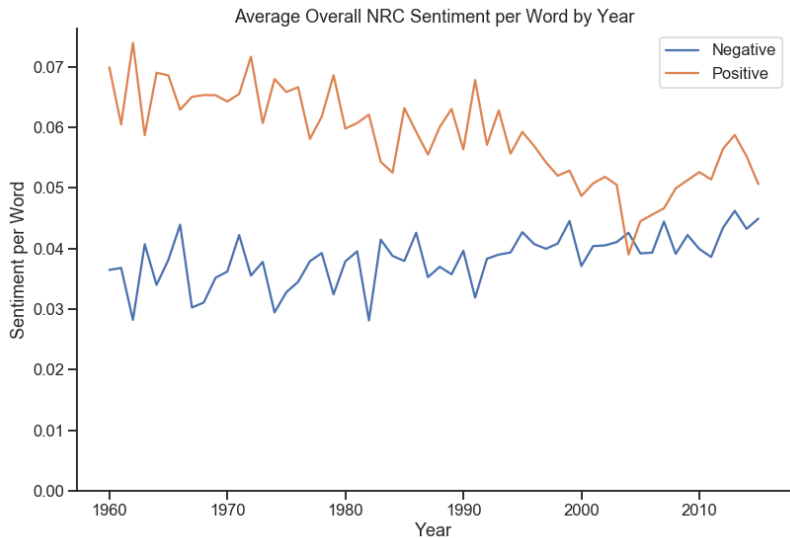
This Is Because There Is Less Joy In Lyrics

Average NRC Sentiment per Word by Year and Type
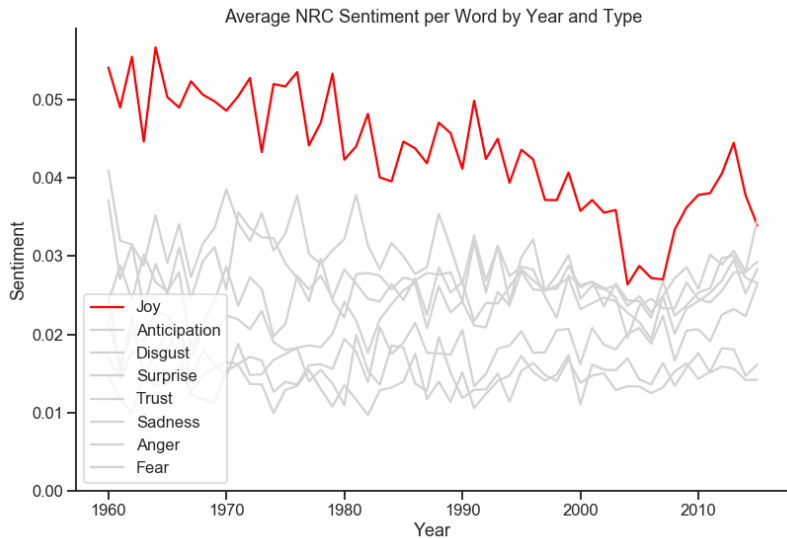
Combined Measurement of Sentiment Goes Down

Average Emotion Index by Year
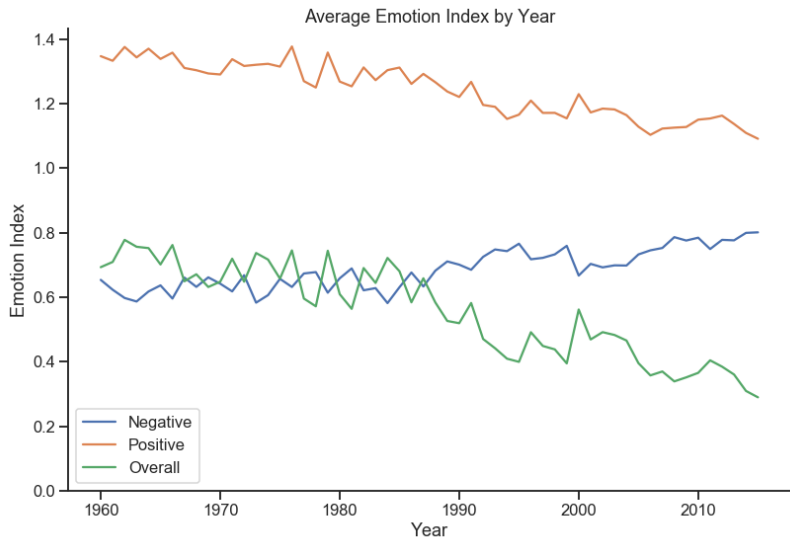
Conclusions

Hits have become:

1. Longer
2. Louder
3. More danceable
4. Less positive (just from audio features)

Lyrics of the hits have become:

1. Longer
2. Faster paced
3. Less positive (especially lack of joy)

Just For Fun

- Most positive song: Hank Ballard and The Midnighters - Finger Poppin' Time, 1960
- Most negative song: Tyga - Rack City, 2012
- Most words: 1536 words, Eminem - Rap God, 2013
- Shortest song: 1 min 36 sec, Dave Clark Five - I Like It Like That, 1974
- Longest song: 26 min, Mike Oldfield - Tubular Bells, 1974

Questions?