Project Work Flow

Upload 'Billboard 100(1950-2018).csv' to Jupyter Notebook

Webscrape genius.com for Lyrics Using Beautiful Soup

Perform Basic Data Analysis

Preprocessing Lyrics Without Punctuation, Lyrics by Phrases, Lyrics by Words with NLTK

Analyze Unique Artists, Artist
Frequency, Unique Phrases,
Average Character and Word
Count

Analyze Frequency of Parts of Speech, Speculative Words, Personal Pronouns, and Lexical Diversity

Explore Freugency of Technology, Swear, Provicative, and Adoration Words

Network Graph and Visualizations

Sentiment Analysis (By Phrases)

Explore Common Words, Direct Count (Without Stopwords), TF-IDF, Topic Modeling

Collaborate and Share Work

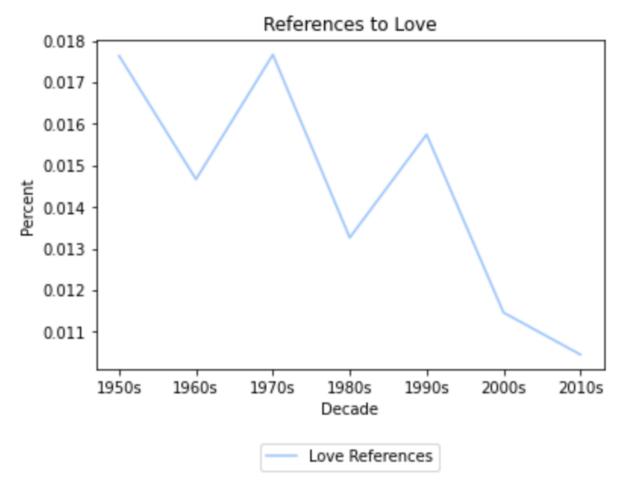
Lyrical Transformation of Billboard

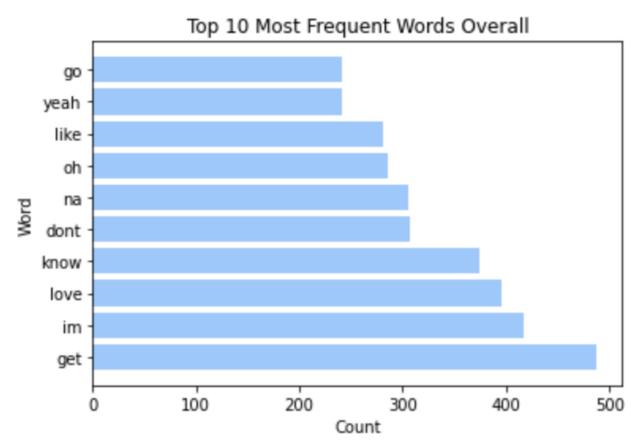
DH 100 Theory and Methods | Alisha Lewis | 6 June 2021

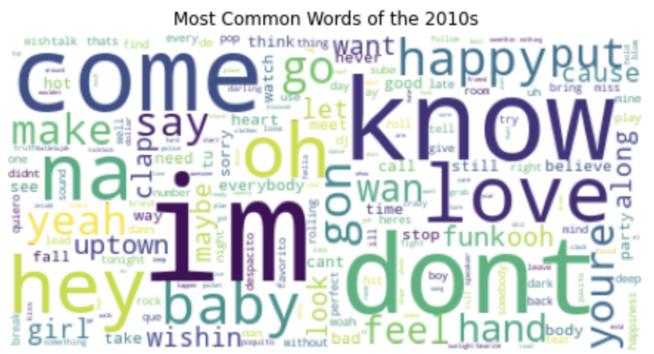
Introduction:

The content of popular music has changed drastically over the years and often highlights the beliefs of our society. Through looking at music lyric trends, values and cultural shifts can be analyzed. While studies have concluded that music has gotten more negative and less about love, this project aims to analyze how much more music mentions technology, sex, pet names, and uses swear words. Society has progressed to be more accepting of these topics in popular culture, and this study hopes consider if this is reflected within popular music lyrics.

Research Question: How have Billboard Hot 100 Lyrics Changed Since 1950 to Reflect Societal Values?







Discussion of Your Results:

The results will hopefully show the frequency of specific themes as well as basic lyric analysis. While the limited number of lyrics can display the graphs to the left, the inclusion of more lyrics will make these results more definitive. The top chart is a line chart that displays the frequency of references to specific themes such as Love, Technology, or Swear Words. Each of these categories have been defined by creating a word bank for association terms. The second chart displays the most common words in a bar chart. The lowest visual is a word cloud that shows the most common words from a specific decade.Each of these visuals are helpful for showing change over time or good displays of frequency. As these graphs are only using lyrics from the top two songs of each year, their accuracy and shape will improve once there are more lyrics.

Billocard HOT 100

Description of Data Set:

This project will use a dataset on <u>Billboard Hot 100 songs from</u> 1950-2018 created by Eunice Chan. This dataset contains the Billboard Hot 100 song title and artist for every year from 1950 till 2018. Eunice manually inputted lyricis for the top two songs of every year, but this project relies on having more data for more detailed analysis. To learm how to web scrape lyrics from genius.com, work from <u>Rishabh Parekh's project</u> will be adapted to this project. Once lyrics have been scraped and added to the Billboard Hot 100 Songs csv, the data set will include the lyrics along with the song title and artist name. While Chan's and Parkekh's work has insipred much of the organization and basic data analysis, this project hopes to expand by looking at specific themes in the lyrics.

Interpreting Your Results:

It could be hypothesized that there will be an increase in references to technology, swear words, pet names, and sex based on our more progressive values in society. With limited data, it has been shown that there are less references to love and more uses of swear words. In the other categories, there have been peaks in other decades so there is no suggested correlation. For technology, there was a peak in the 80s, and is now on the rise again. When there is more lyrical data, this project should be able to conclude of this is just by chance or the more common use of phones starting in the 80s.

Conclusions:

By the end of this project, an argument can be made about the topical matter of lyrics from the 1950s to today. Future work could include expanding the dictionaries to include more words of association. Additionally, work could be done on looking at the most common words each year and determining a category for those.

Work Cited:

More Negative Lyrics
Chan, Billboard Hot 100
Parekh, DH Project