

# Lyrical Transformation of Billboard HOT 100

DH 100 Theory and Methods | Alisha Lewis | 27 June 2021

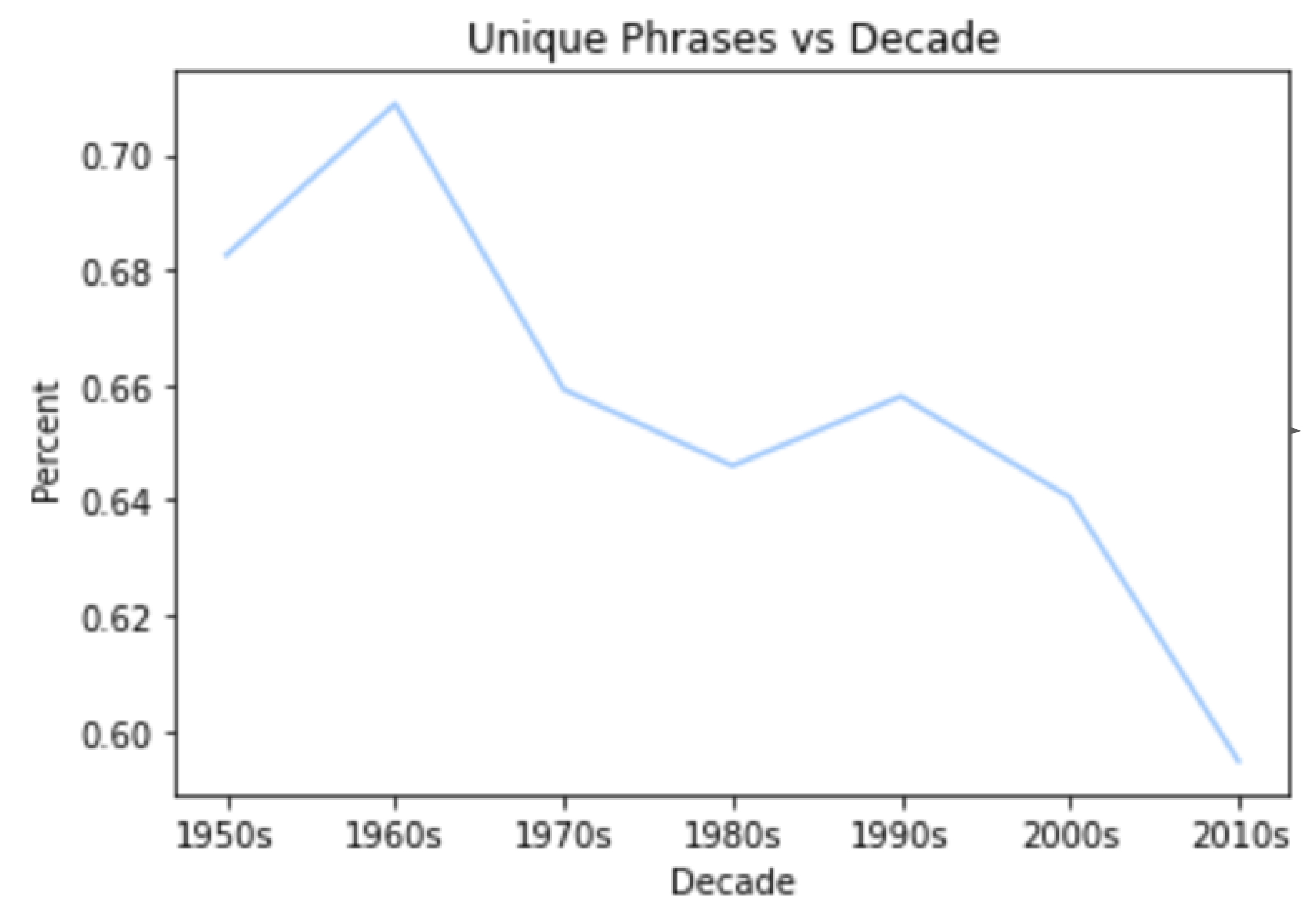
## Introduction:

The content of popular music has changed drastically over the years and often highlights the beliefs of our society. Through looking at American 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 and swear words, along with more general analysis. America 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?

## Description of Data Set:

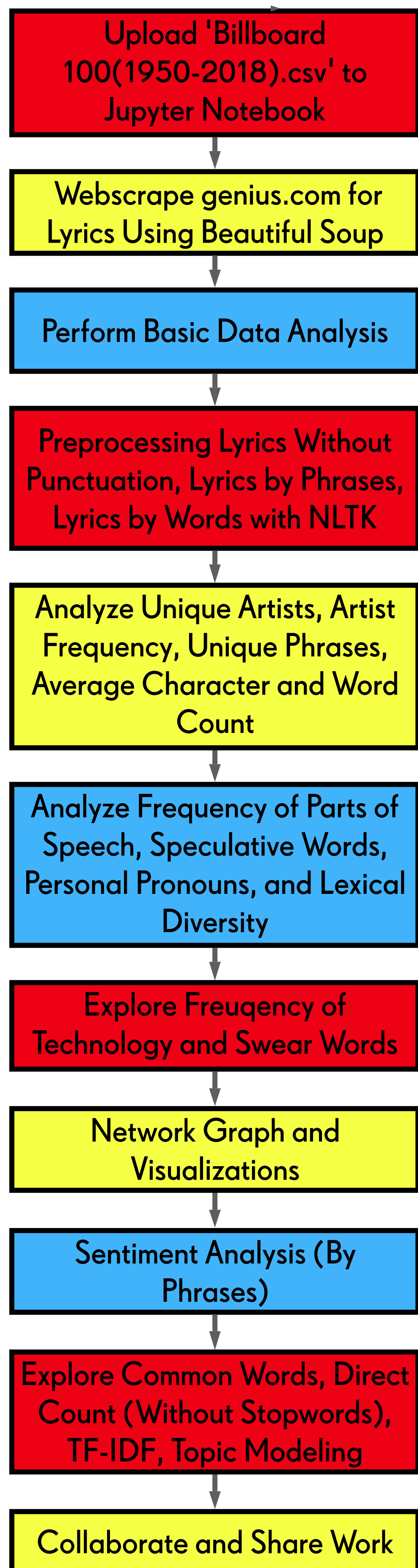
This project will use a dataset on [Billboard Hot 100 songs from 1950-2018](#) created by Eunice Chan. This dataset contains the Billboard Hot 100 song title and artist for every year from 1950 till 2018. Eunice manually inputted lyrics for the top two songs of every year creating 136 data points, but this project relies on having more data for more detailed analysis. To learn how to web scrape lyrics from genius.com, work from [Rishabh Parekh's project](#) will be adapted to this project. Using Parekh's technique, the lyrics for the songs in Chan's csv file were scraped. Some lyrics were unable to be found so there are 2,540 song lyrics in the expanded data set. This dataset includes the lyrics along with the artist, song, and year. While Chan's and Parkekh's work has inspired much of the organization and basic data analysis, this project hopes to expand by looking at specific themes in a wider dataset of lyrics such as technology and swear words. With a much larger dataset, this project is able to produce more reliable results that show better trends about lyrics have changed generally. To analyze specific shifts in references to technology and swear words, a dictionary was manually created to track how many times those words appear in the lyrics. Using data analysis, NLTK, and visualizations, this project will highlight how Billboard Hot 100 lyrics have changed since 1950, reflecting a shift in societal values.



## Interpreting Results:

In the graph "Unique Phrases vs Decade," there appears to be a steep decline in the amount of unique phrases since 1950. While there was a slight bump in 1990, there is an overall downward trend. In the words clouds for the 1950s and 2010s, there are both similarities and differences. While words such as "love" and "baby" are not the most popular words by 2010, they are still common enough to be featured in the word cloud. There are words that disappeared completely from the most common words since 1950, such as "rock" and "kiss." Additionally, words have appeared in the 2010s that were not present before like "ain't." In the "Technology References" graph there are two spikes in the 1970s and 2010s. The 1990s appear to have the lowest amount of technology references. Looking at the "Use of Swear Words" graph, there is a steep incline of swear word usage starting in the 1990s. It seems to have leveled off slightly in the 2010s, but there is still an increase. Prior to that, there was a slight increase in the 1970s.

## Project Work Flow



## Discussion of Results:

It could be hypothesized that the trends these graphs are correlated to shifts in our societal values. As our society has become more progressive and open, those shifts have been reflected in our popular music. A possible explanation for a decrease in unique phrases is the rise in rap music. This genre uses shorter, more repetitive phrases and began to be popular around the 1970s. When considering shifts in the most common words, it can be assumed that this is due to changes in our language. For example, it would make sense that "ain't" did not appear in the 1950s most common word cloud because it was not apart of the slang at that time. The lyrics seem to be less about love, suggesting that the US has become less romantic and could be focused on other topics such as sex. Furthermore, words like "rock" may be less common because of a shift in genre popularity. Rock was the most popular genre in the 1950s, but hip hop popularity has risen in 2010. When considering references to technology, there is a spike in the 1970s, which is when the first personal computers were created. There is another spike in the 2010s, which could be explained by the more common use of cell phones. It is unusual that there were more references to technology in the 1950s and 1960s than there were in the 1990s and 2000s. It is possible this could be due to the selection of words in the technology dictionary. The use of swear words starkly increased in the 1980s, which could be due to the popularity of rap starting in the 1970s. Swearing also suggests a more casual societal standards.

