

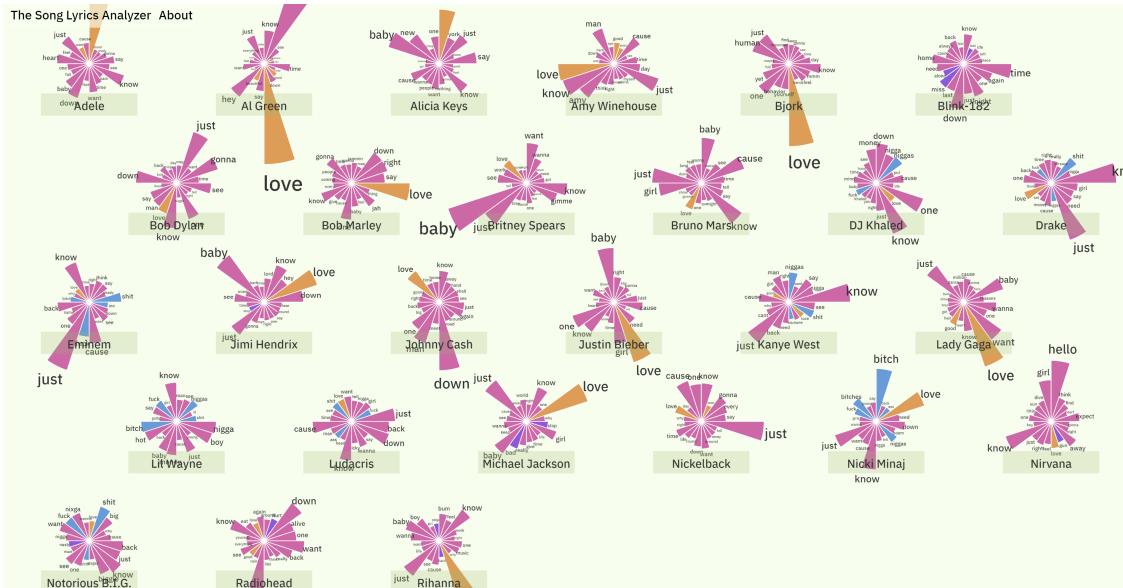
Final Written Report

Song Lyrics Analyzer

Valentine Sénégas

CART 451 - Networked Media Studio
Presented to Sabine Rosenberg
December 9th 2022

Project Description



This project analyzes song's lyrics from major artists from the 20th and the 21st century. A graphic representation shows the the most used words for each artist.

Technologies

I used **Natural**, a **natural language processing**, tool to process the lyrics of the different songs.

I used the **tokenizer** to separate each word.

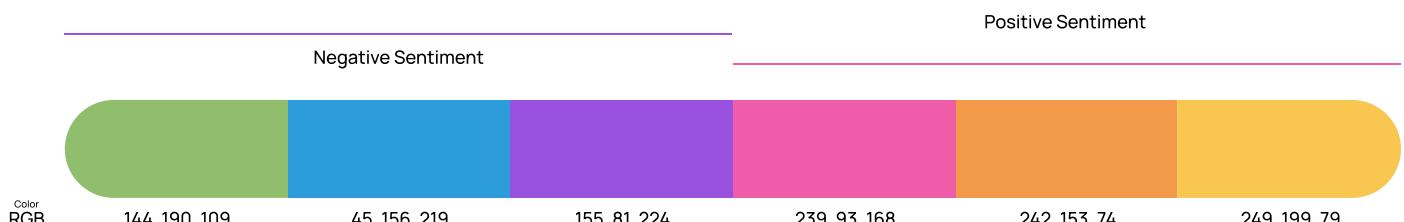
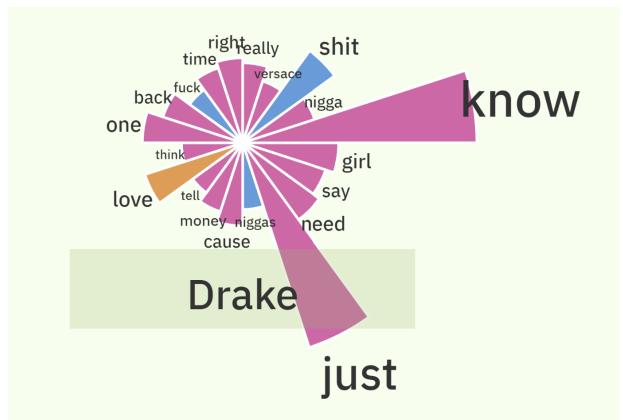
I used the Term Frequency–Inverse Document Frequency (**tf-idf**) to know which words are the most frequent ones for each artist. The 20 most frequent words were used to create the stars.

To create the “Stars”, which are the visual representation of song lyrics, I used several technologies. The arcs are drawn using **p5.js**.

I used **Sentiment Analysis**, an algorithm of **Natural**, to determine the sentiment associated with each word. This sentiment determined the color of each arc.

Features

“Stars” are the main feature of my project. Each star represents the diversity of an artist’s lyrics. A star contains twenty arcs, which all represent a single word. These are the twenty most used words of the artist. The longer it is, the more times the word was found in the database for this artist.



Every arc has a color, and although it might seem random, it is not. Words associated with negative feelings have cold colors and words associated with a positive sentiment have warm colors. This color chart is an additional feature in the about page. It is there to help the user understand the meaning of each color.

A feature that I would have liked to implement is letting users add a new artist. It would have been a really nice feature to have, as it would become collaborative.

Add your own artist!

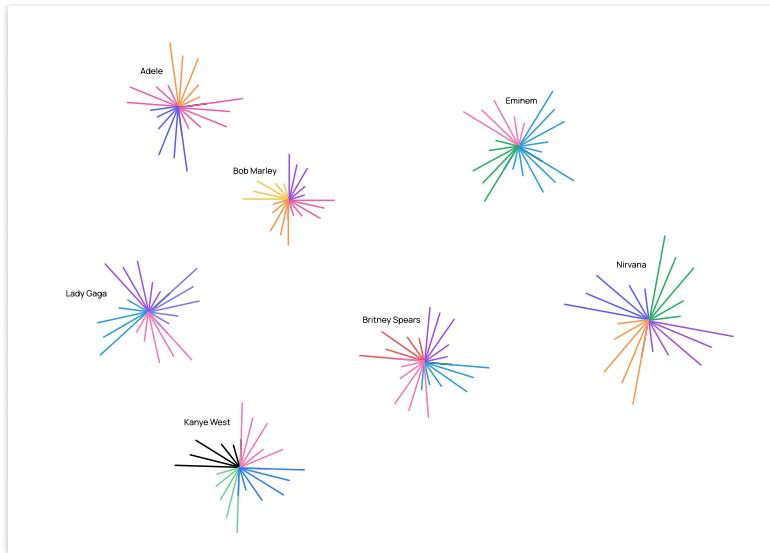
Name of the artist

Enter the lyrics of their songs.
Do not input [] {} or ;

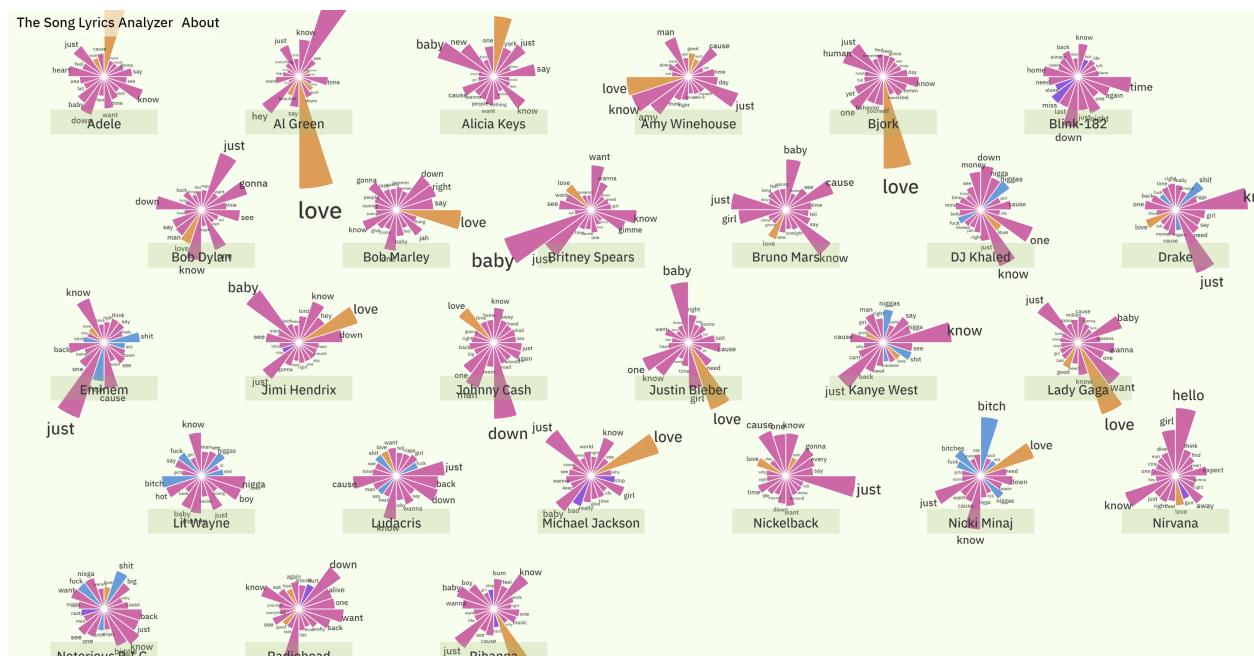
Send

How the project implements the intended idea

Below is a design of the intended creation and a screenshot of the final project. I believe that I have successfully achieved my goal.

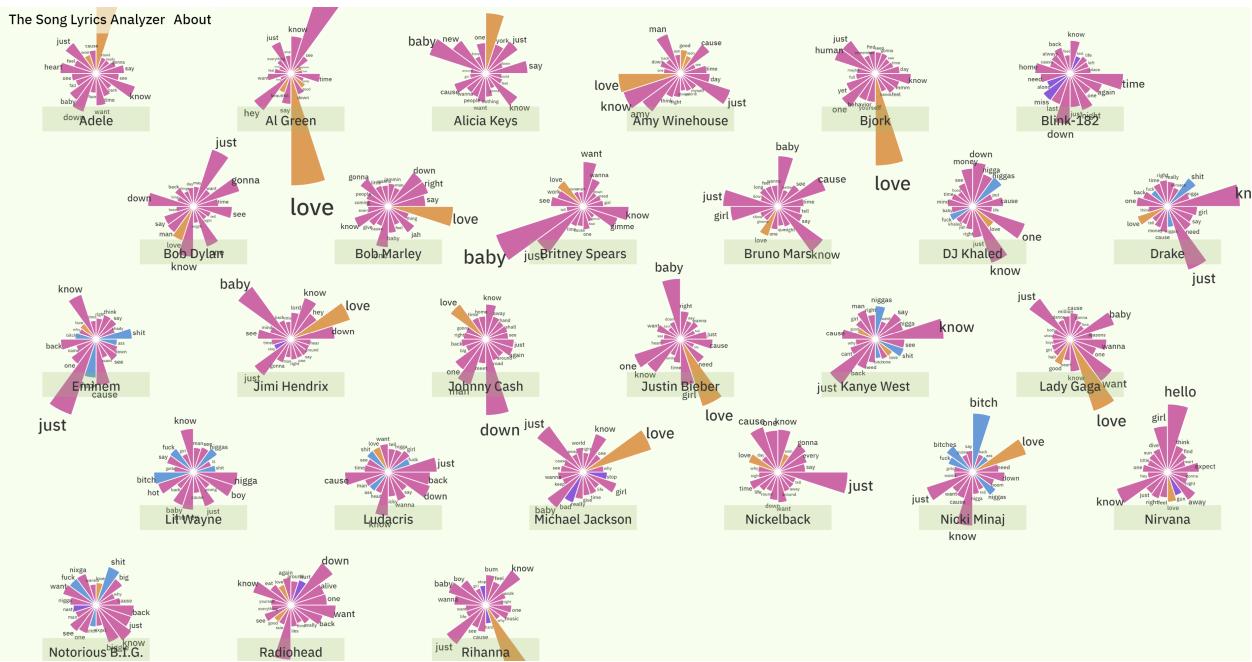


Intended creation

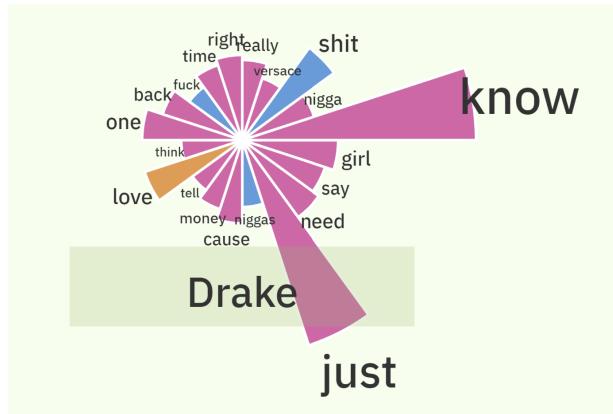


Final project

Screenshots



Overall project



Example of a star, zoomed in

Screenshots

About the project

This project analyzes song lyrics from more than 40 major artists from the 20th and the 21st century. The main goal is to compare word similarities or differences between all the artists.

A graphic representation displays a set of words for each artist, which will be linked to others. I call these visual representations "stars". Each star is made out of 20 arcs, and each one of them represents a word.

Each word is associated with a color. The sentiment associated with the word will determine its color.



One of the big observations that can be made is that most of the artists have the word "[love](#)" as their most spoken word! Another thing that we can notice is that the majority of the words are represented as pink, which means that they are positive words.

If we look into more detail and compare the genres, we can clearly see that pop artists use very positive words because they are represented as orange and pink arcs.

On the other hand, rap artists tend to use more negative words, as we can see more blue and purple.

Finally, Blink-182, the only punk rock band, was the one that had no orange and three arcs of purple, which shows that the words are mostly negative.

I was expecting to see a wider diversity of color.

About page explaining the color chart

Inspiration from in-class readings

Hito Steyerl

Hito Steyerl's text highlights the importance of putting data in context. She starts the text by showing an abstract image revealed by Edward Snowden. Taken out of context, this image does mean anything: "Snowden's image of noise could stand in for a more general human inability to perceive technical signals unless they are processed and translated accordingly". In a similar way, my project extracts words from songs, exactly as they were written, and try to make analogies and conclusions from these individual words. My project does not consider groups of words or expressions. Although the data will be accurate strictly considering the words, the meaning behind the songs is not reflected in the end result.

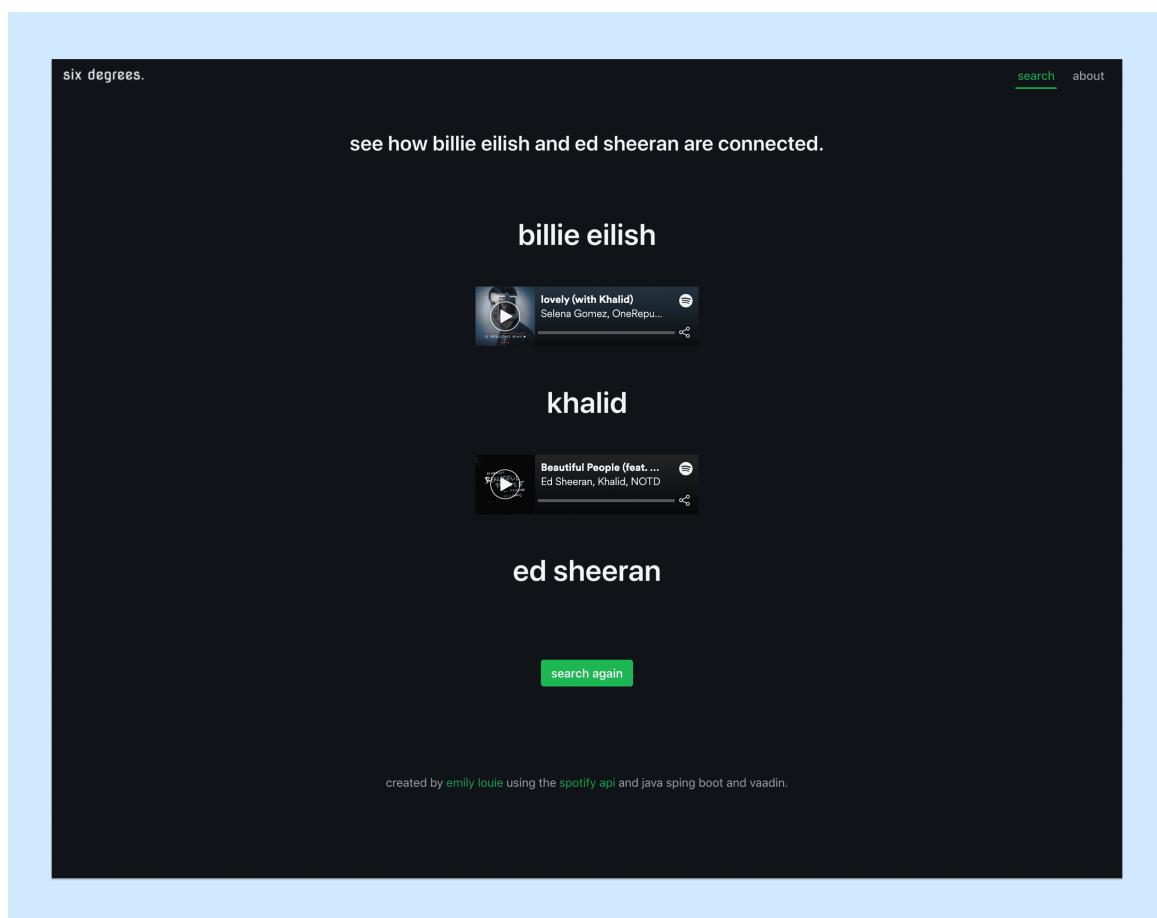
Mimi Onuoha

At first, I thought that my dataset could not be biased because the lyrics of the songs can easily be verified. Then, I realized that the artists and songs in this dataset are only a small fraction of all pop music. My project demonstrates that pop artists use a lot of swear words. The project will not be able to make sense of all the words together and analyze them individually, which could distort the intended meaning behind the songs.

Other Inspiration

Six Degrees of Spotify

A very interesting project made using a lot of data is Six Degrees of Spotify by Emily Louie. In 2020, she created a website that relates the theory of six degrees of separation to music. This theory is that every person in the world is connected to anyone in less than 6 connections. Using the Spotify API, she was able to determine how two artists were connected to each other by looking at their collaborations on songs.



Other Inspiration

Music-map.com

Music-map.com is a website illustrating the connections between different music artists. The user inputs the name of an artist, and a representation appears, showing similar music artists. The user can then click on each artist to see artists similar to this one. I used a similar system to visualize my data.

