RISE RESEARCH INNOVATION SCHOLARSHIP ENTREPRENEURSHIP

Mentor Name. : John Rachlin
Subject/Category : Data Science
Class : Advanced Programming with Data (DS3500)
Group # : 41

Northeastern University

Analyzing Beatles Song Lyrics Over Time

Sophie Sawyers; John Rachlin

BACKGROUND

Figure 1: Background. The Beatles are widely considered one of the greatest bands of all time, capturing audiences of all ages for the last sixty years. However, The Beatles' success, style, and image are anything but formulaic, reflecting the experimental mindset and departure from societal norms that dominated the spirit of the 1960s at the height of their fame. This project aims to examine Beatles song lyrics across albums released between 1962 and 1970 to identify and characterize these stylistic changes to better understand The Beatles' musical evolution.

METHODS

lyricool.py

A reusable library for lyric analysis and comparison

lyricool_app.py

Apply text analysis library to multiple sets of lyrics

lyricool_parsers.py

A custom parser to pre-process and parse lyrics from AZLyrics.com

Libraries: NumPy, matplotlib, pandas, TextBlob, HUGGING FACE, BeautifulSoup

Figure 2: Lyricool Architecture. The Lyricool framework consists of three files: lyricool.py where the library is created, lyricool_app.py where the library is applied to song lyrics, and lyricool_parsers.py where custom parsers are defined for parsing from domain-specific files (in this case, AZLyrics.com). Multiple libraries were used to develop the framework, including NumPy, matplotlib, pandas, TextBlob, huggingface, and BeautifulSoup.

PRODUCTS

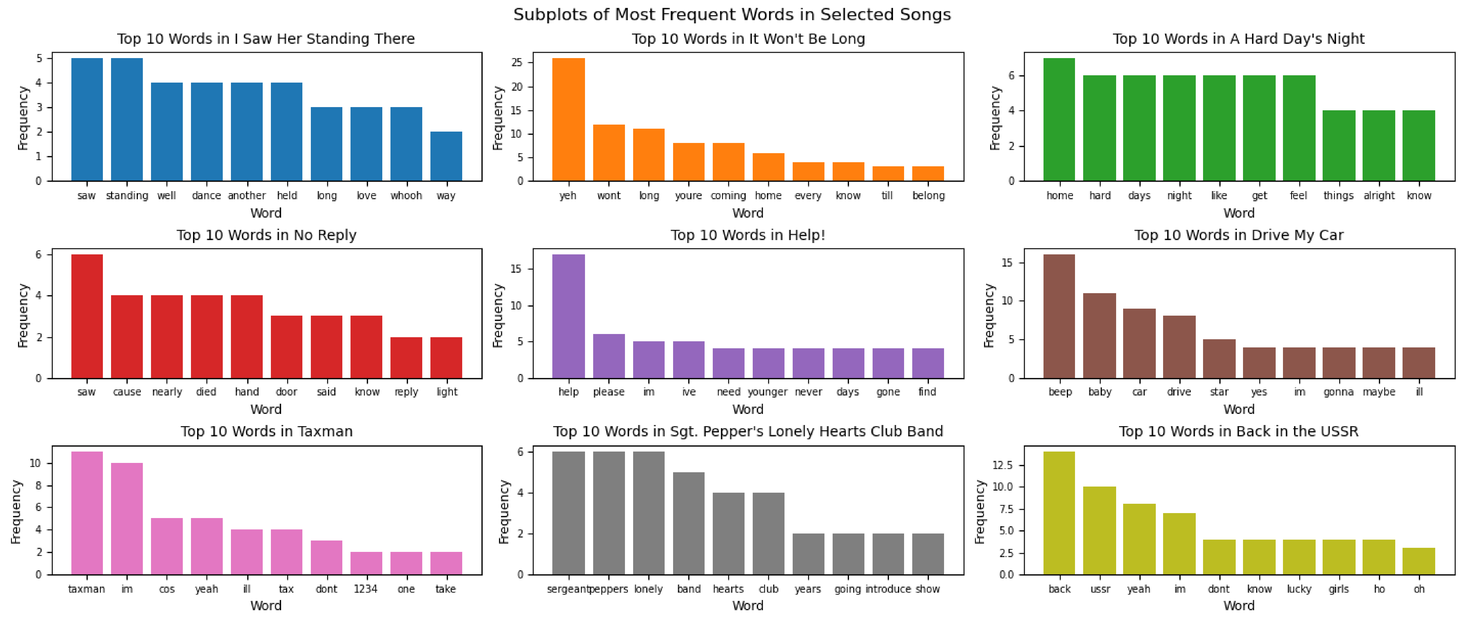
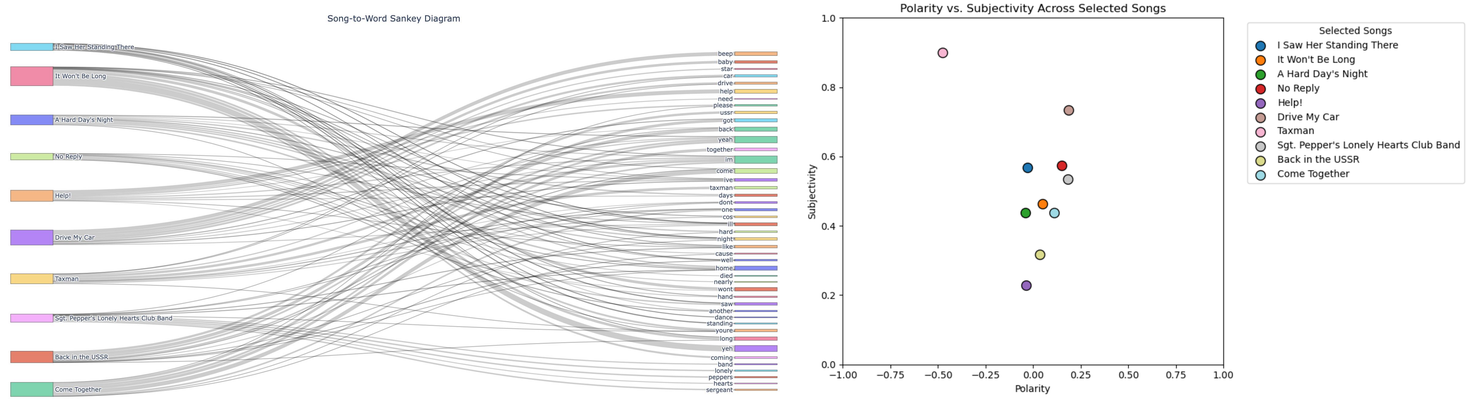


Figure 3: Lyricool_app.py Output. When executed, lyricool_app.py produces three visualizations, a song-to-word Sankey diagram (bottom left) which maps each set of lyrics to words (either *k* most common or a user-entered list), a figure comparing word frequencies of each set of lyrics using bar plot subplots (top), and a scatterplot (bottom right) comparing polarity vs. subjectivity scores across selected sets of lyrics/songs. In this case, we are comparing the first songs of the 10 albums between "Please Please Me" and "Abbey Road" in chronological order of release.



FINDINGS

The analysis of Beatles lyrics reveals thematic shifts across their discography. Early albums ("Please Please Me" to "Beatles for Sale") center on love and fleeting heartbreak, reflecting their pop origins. Mid-phase albums ("Help!" to "Revolver") explore deeper themes of self-doubt, political critique (e.g., "Taxman"), and existential questioning. Late albums ("Sgt. Pepper's Lonely Hearts Club Band" to "Abbey Road") showcase experimental approaches influenced by psychedelics and innovative production techniques, signaling their artistic evolution. The song-to-word Sankey diagram highlights word overlaps across albums, with "It Won't Be Long" and "Sgt. Pepper's Lonely Hearts Club Band" sharing the most words with others (second and eighth album, respectively). Within the terms most repeated, "home" (5) and "night" (4) particularly show how The Beatles evolve from romantic metaphors in early albums to literal references in later ones, reflecting a shift in tone and maturity. The polarity-subjectivity scatterplot shows most songs as sentimentally neutral (0 polarity) and balanced in subjectivity (0.5). Notable outliers like "Taxman" (negative and subjective), "Drive My Car" (more subjective), "Help!", or even "Back in the USSR" (more objective) reveal thematic and emotional diversity over time as these albums differ in phases, although all earlier albums seem to be consistently balanced in polarity and subjectivity. These findings provide a data-driven perspective on The Beatles' lyrical evolution, linking their themes to broader societal shifts of the 1960s, including changing views on love, identity, and politics. If given more time, I would like to analyze all songs by album to analyze how tone can change within a given album. I would also like to add more albums to the analysis, such as "Magical Mystery Tour" or "Let It Be", and compare Beatles lyrics to lyrics of other artists from the 1960s to contextualize The Beatles' audience appeal at their peak.