# THE SOUND OF MUSIC: A SPOTIFY-BASED ANALYSIS OF GLOBAL MUSICAL EVOLUTION

**Exploring the dynamic world of music, this project analyzes Spotify data to reveal global musical trends and transformations. We combine data-driven insights and interactive visualizations to highlight key artists and evolving styles, offering a comprehensive view of the musical landscape's evolution through the lens of Spotify.**


**CSE 6242 Team #134**
Jacob Marquez, Saroj Thapa, Valentina Zapata, Xuesi Ruan, Zaid Hasan


## INTRODUCTION

Current music analysis is narrow, lacking a holistic view across regions and genres. This hinders aspiring artists and industry stakeholders as they lack essential insights into the factors contributing to the success of diverse musical genres on a worldwide scale.

Our project seeks to address this gap by providing a more nuanced understanding of the intricate patterns shaping the music industry.

## APPROACHES

**Anomaly Detection:**
We utilized the Isolation Forest algorithm to identify emerging trends in musicality and the artists responsible for them .

**Visualizations:**

1. Interactive D3.js line chart that dynamically showcases average song characteristics by genre across decades, offering a user-friendly exploration of musical trends.
2. Interactive map built in ArcGIS Online that displays song origins over time by genre. Users can view songs' metadata and music attributes by clicking on the map dots.

## DATA

The foundational dataset for this project comes from Spotify and contains details of songs' musicality (e.g. danceability, happiness, etc.), popularity, and credits.

We enriched the Spotify data with genre and geographical information via publicly accessible APIs and Python Geo libraries. The enrichment led to a reduction in the overall dataset to ~60,000 records, enough for a meaningful analysis.

## EXPERIMENTS & RESULTS

**EDA:**

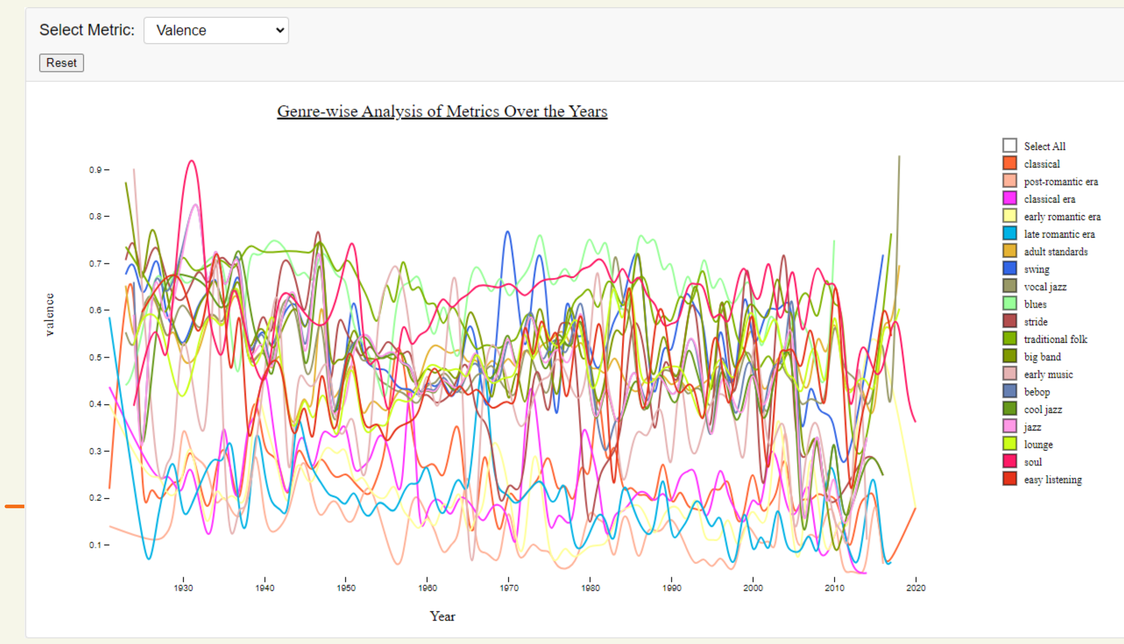
- Observed a consistent 2% decrease in song duration post-2010
- Identified a decrease in the rate of change in musical characteristics since the 1960s
- Found a significant positive correlation between song popularity and energy/loudness (coefficients of 0.49 and 0.46 respectively)
- Discovered a negative correlation (-0.57) between popularity and acousticness

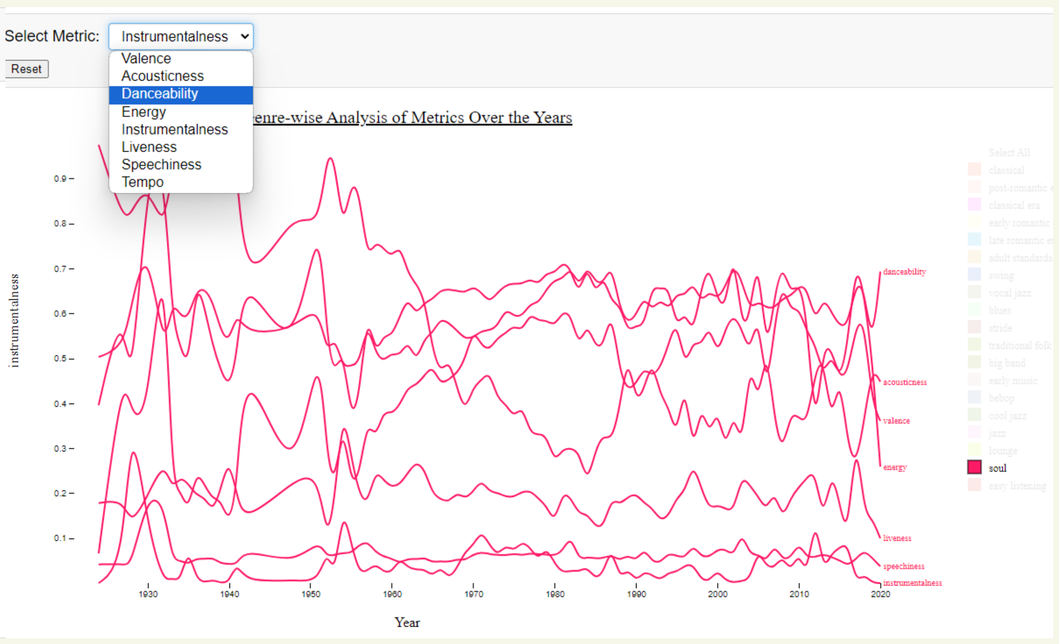
**Anomalies:**

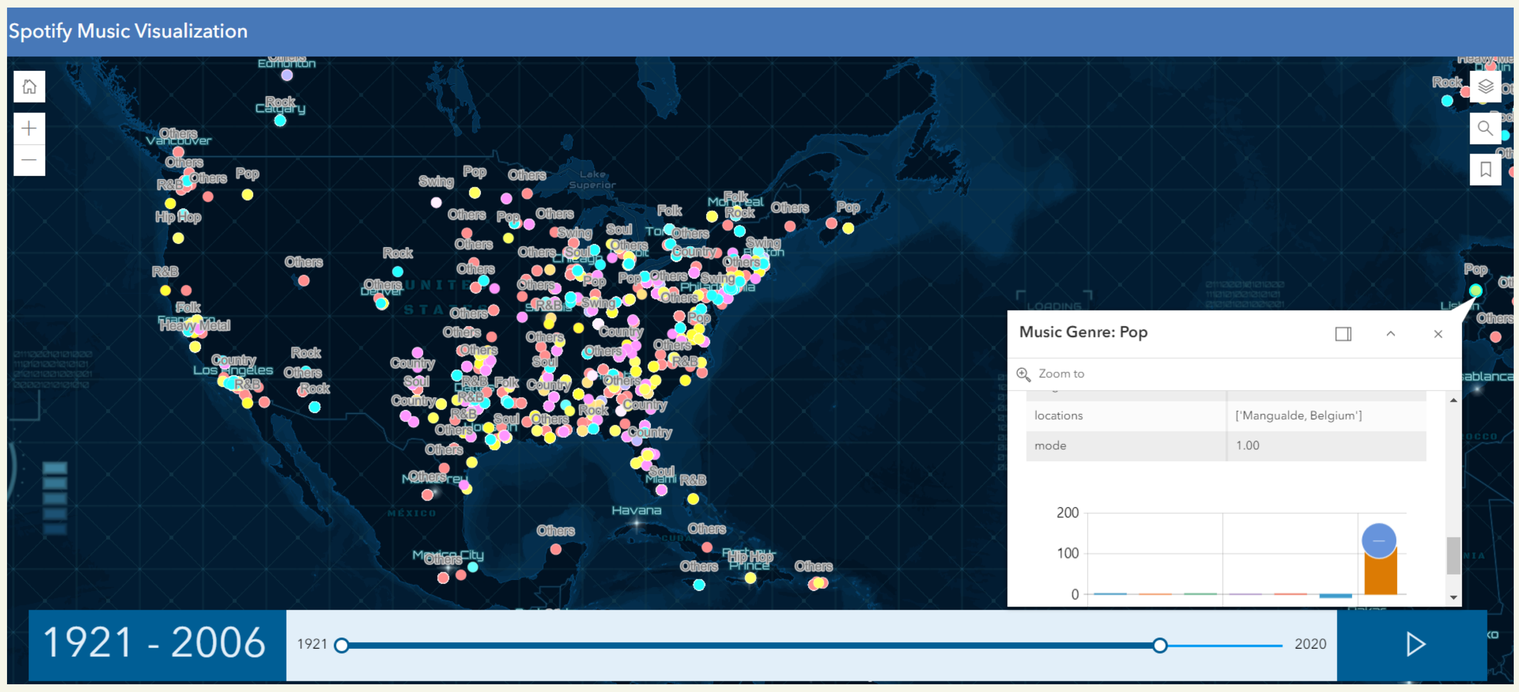
- Anomalies represent older recording eras with trends such as higher acousticness, instrumentalness, and possibly live recording styles, along with longer durations.
- Identified influential artists like Tommy Dorsey, Benny Goodman, Lata Mangeshkar, and Louis Armstrong, who significantly shaped genres like Big Band, Swing, Classical, and Jazz during the early to mid-20th century

**Line Chart Visualization:**

- Revealed consistent patterns in tempo, speechiness, and liveness across most genres, suggesting a shared historical impact.
- Valence, instrumentalness, and danceability exhibit wide-ranging values across genres, undergoing significant changes over time.
- Energy levels in classical genres trend downward overtime, while modern genres trend upwards (opposite behavior for acousticness)




**Map Visualization:**

- Pre-World War II, musical origins were concentrated in the Eastern United States and Europe across all genres.
- Post-World War II exhibits a significant globalization trend, with artists from South and Central America, Africa, and Asia.
- The 1970s mark a surge in musical production especially in genres like pop, R&B, heavy metal, and country.