

BILLBOARD TOP HIT SONGS ANALYSIS & VISUALIZATION

Group Project I - Team 10

Nguyen Tung Lam - V202100571

Nguyen Nhat Minh - V202100570

Nguyen Truong Tung - V202100542



DATASET OVERVIEW

- The dataset consists of the weekly Billboard Hot 100 chart spanning from January 4, 1960, to April 2, 2022, containing a total of 324,787 records.
- The data is filtered based on the number of weeks each song remained on the chart (as an indicator of popularity). There are 3150 records as a sub-dataset.
- Retrieved the top 50 songs with respect to their weeks on chart.

billboard



Spotify®

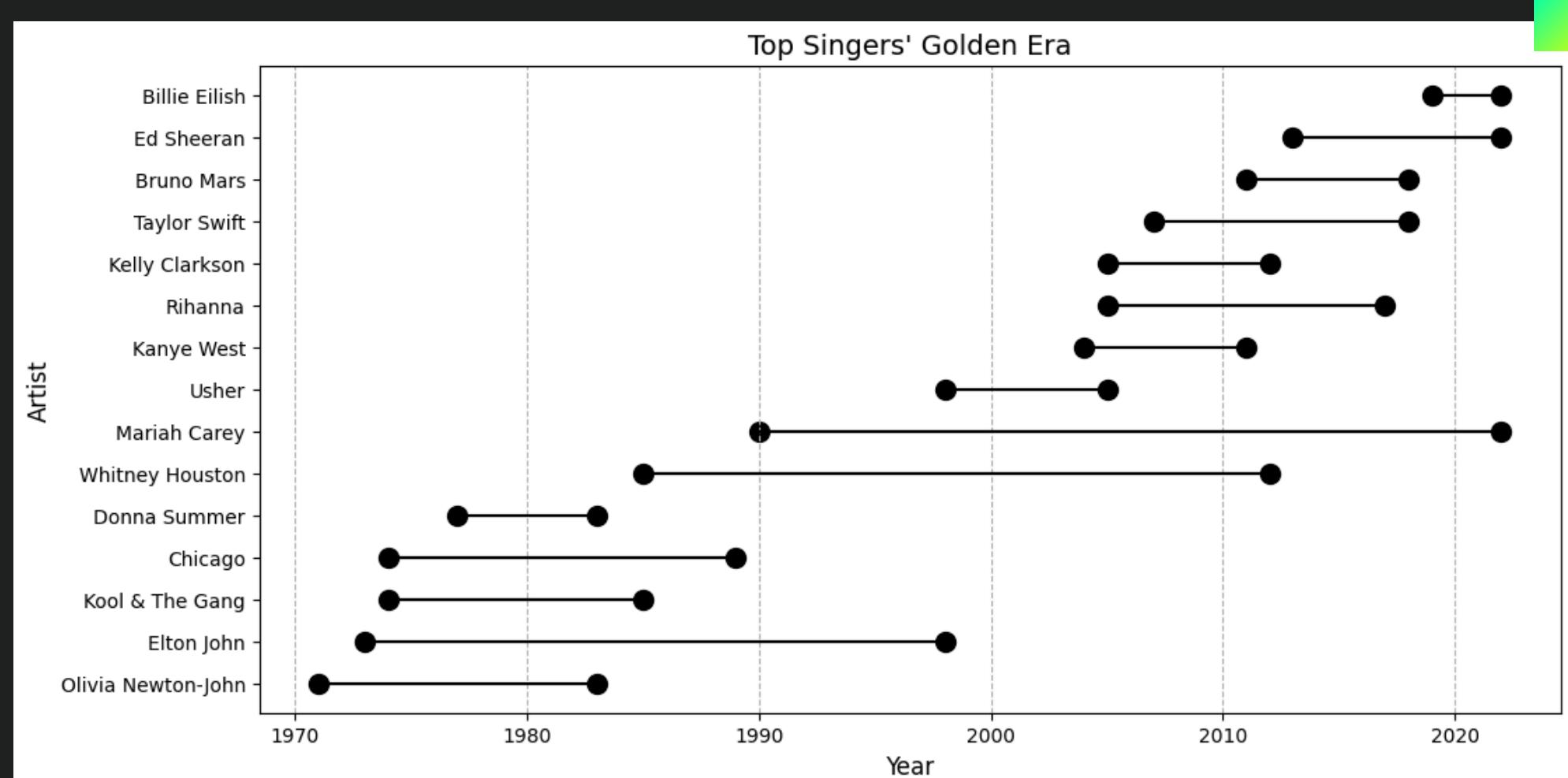
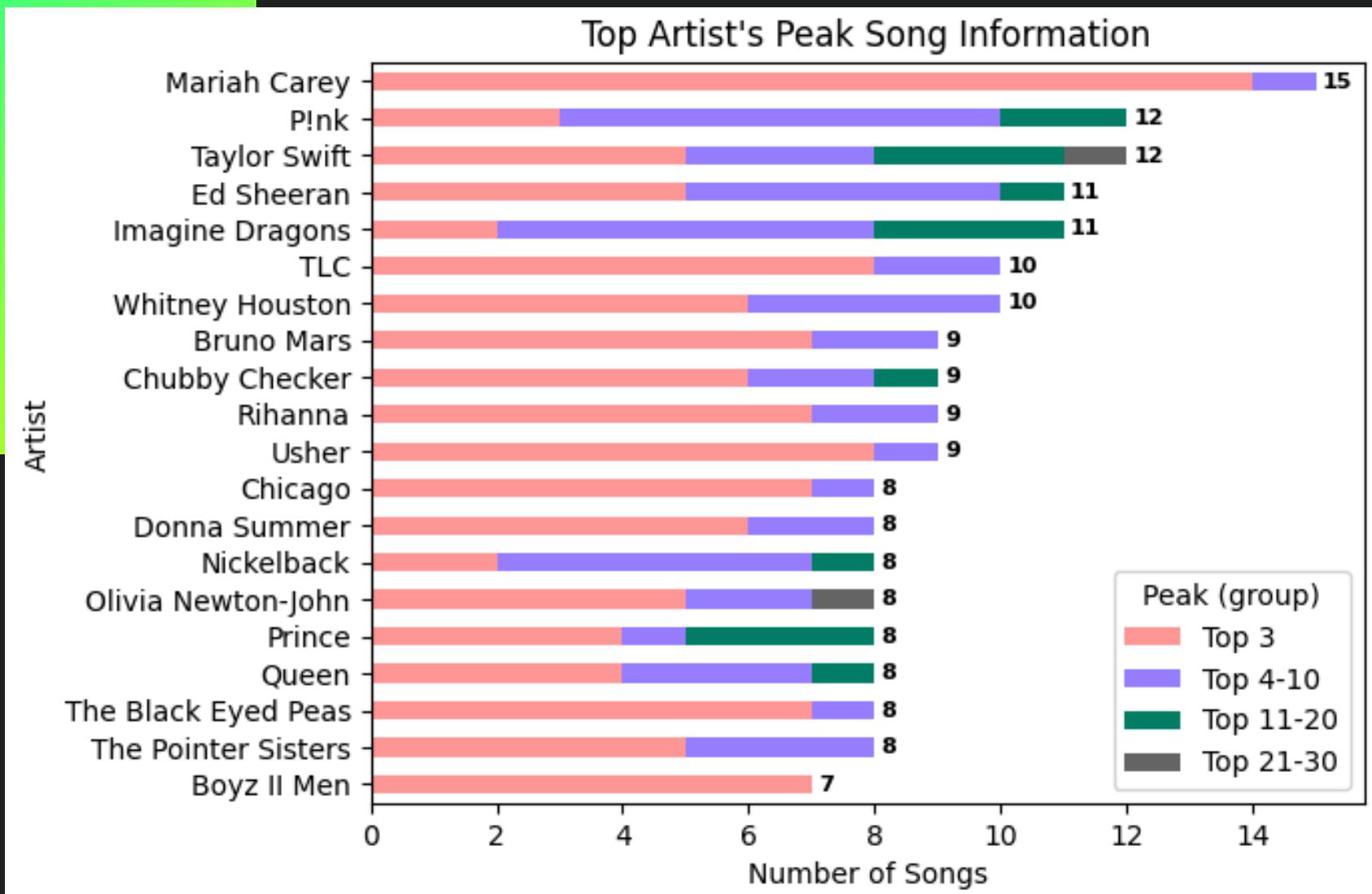


DATASET OVERVIEW

- Retrieve the data of each song's audio features such as danceability, acousticness, loudness, key (pitch class), artist's genre, tempo, valence (positiveness of the song), etc from Spotify.
- For each record, we have 21 attributes in total.



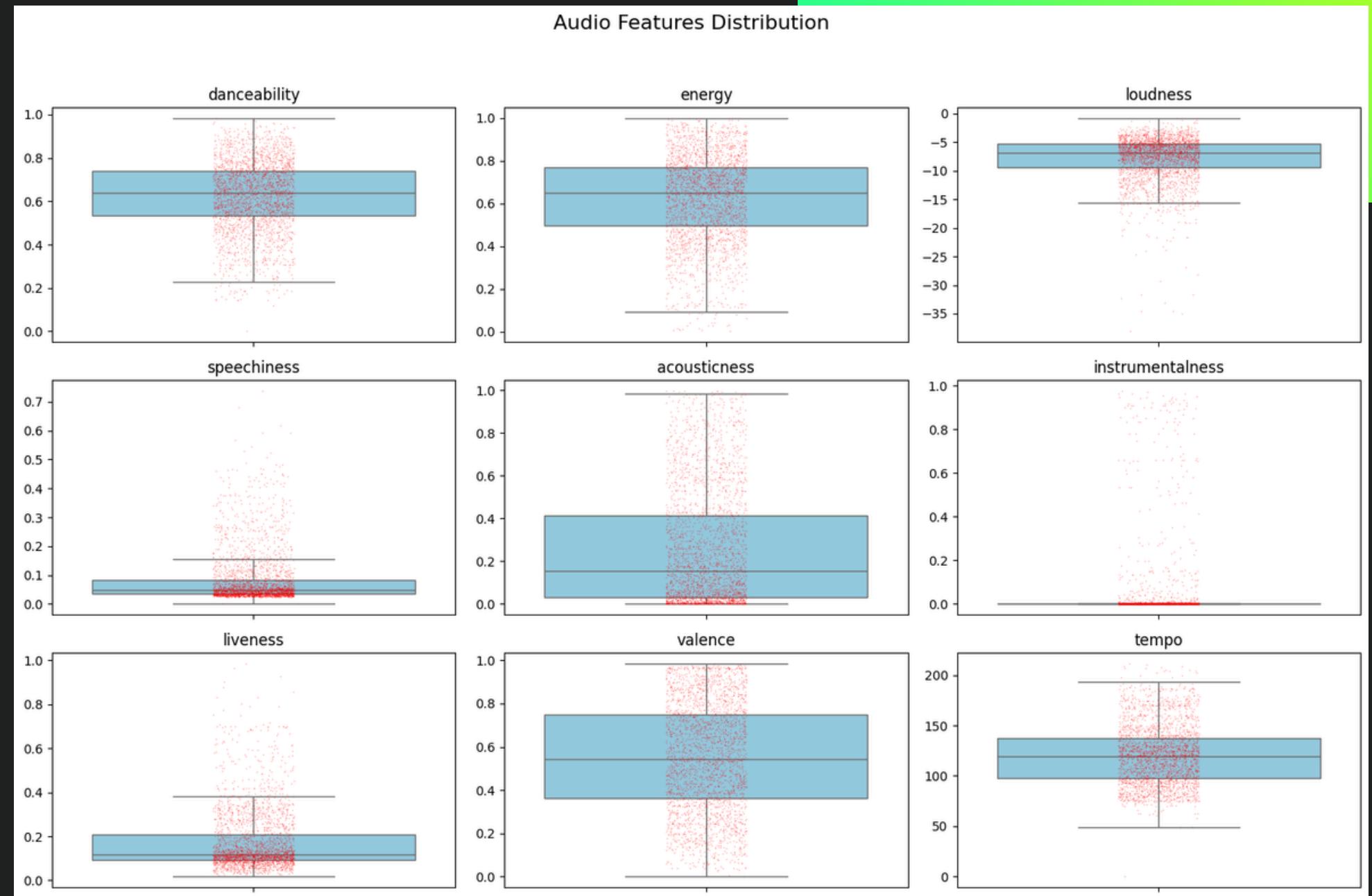
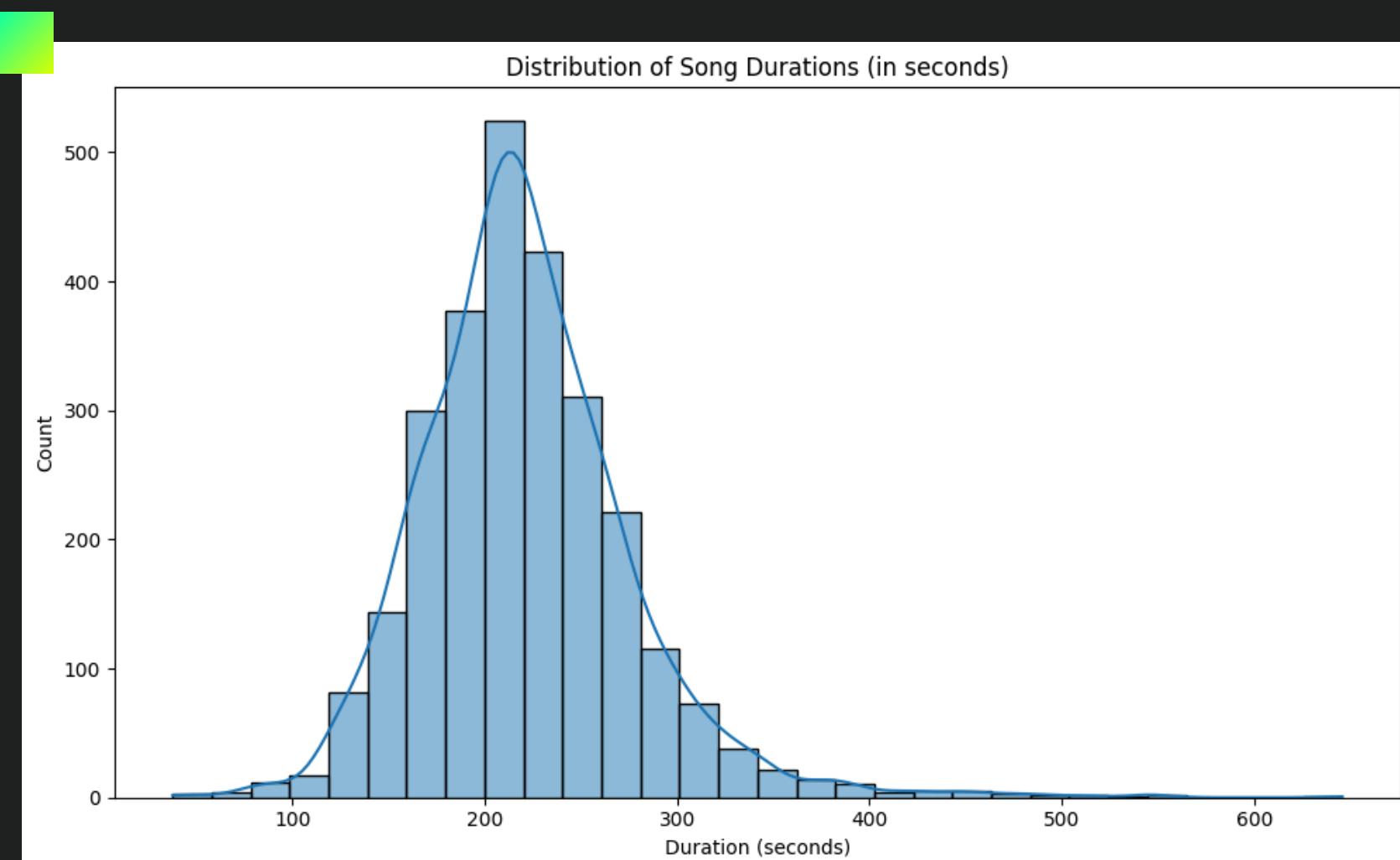
Column	Description
date	The date when the song appeared on the Billboard chart.
title	The title of the song.
artist	The name of the performing artist.
year	The year of the song's chart performance.
rank	The song's ranking on the chart for the given date.
last_week	The song's ranking in the previous week.
peak	The highest rank the song achieved on the Billboard chart.
weeks	The total number of weeks the song remained on the chart.
danceability	A measure of how suitable a track is for dancing (0 to 1).
energy	A measure of intensity and activity (0 to 1).
key	The musical key in which the song is played (0-11, corresponding to C to B).
loudness	The overall volume of the track in decibels (dB).
mode	The modality of the song (0 = Minor, 1 = Major).
speechiness	The presence of spoken words in a track (0 to 1).
acousticness	The likelihood of the track being acoustic (0 to 1).
instrumentalness	The likelihood of the track being purely instrumental (0 to 1).
liveness	Detects the presence of a live audience in the recording (0 to 1).
valence	The musical positiveness of the track (0 to 1).
tempo	The estimated beats per minute (BPM) of the track.
duration_ms	The length of the song in milliseconds.
genres	A list of genres associated with the song.
genre_encoding	A categorical encoding of the song's genre for classification tasks.

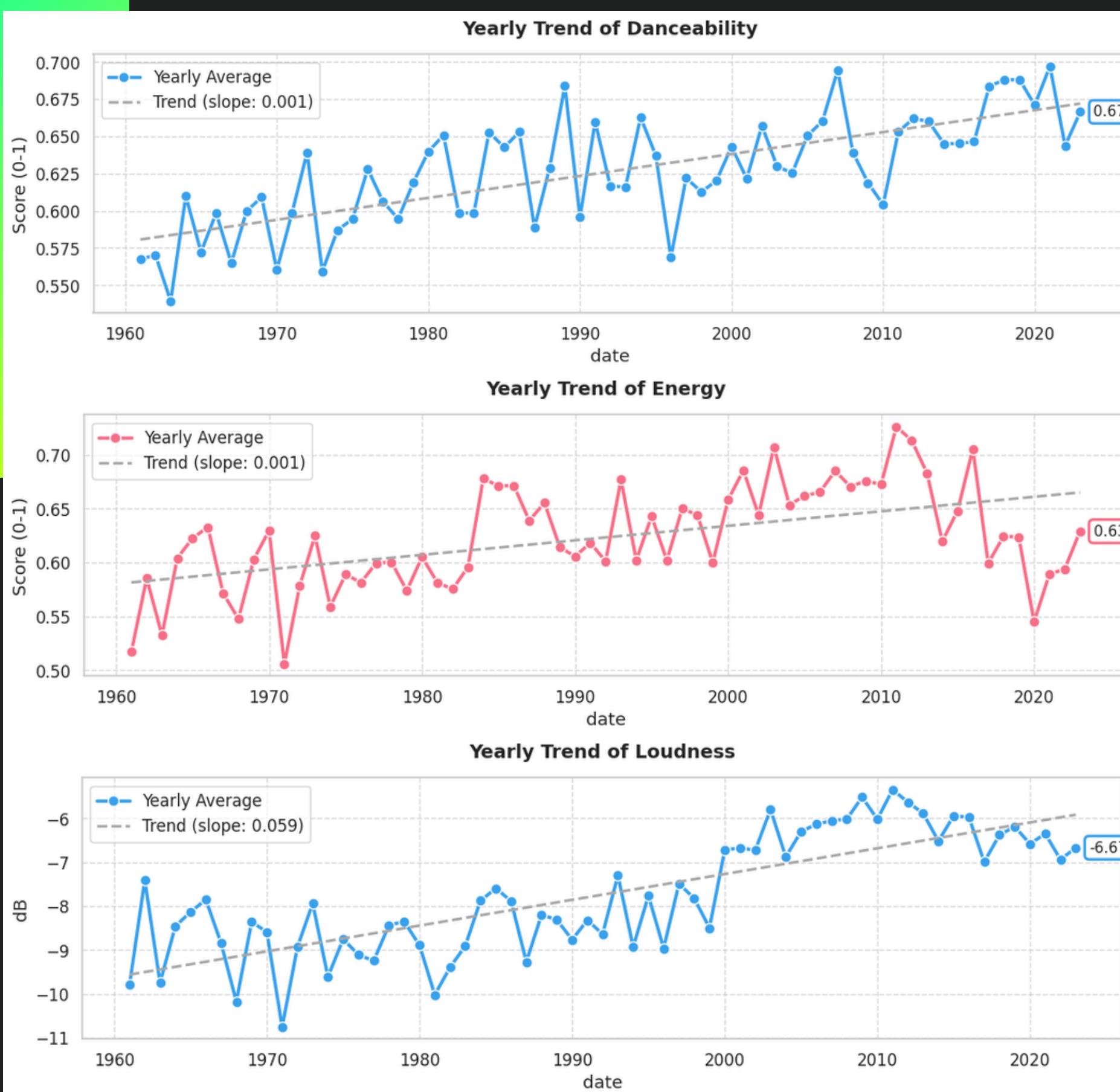




WHAT ARE THE KEY AUDIO FEATURES THAT DEFINE TOP- CHARTING SONGS OVER THE PAST DECADE?







MUSIC TRENDS OVER TIME (PT. 1)

01

Danceability, energy, and loudness on averages all increases YOY

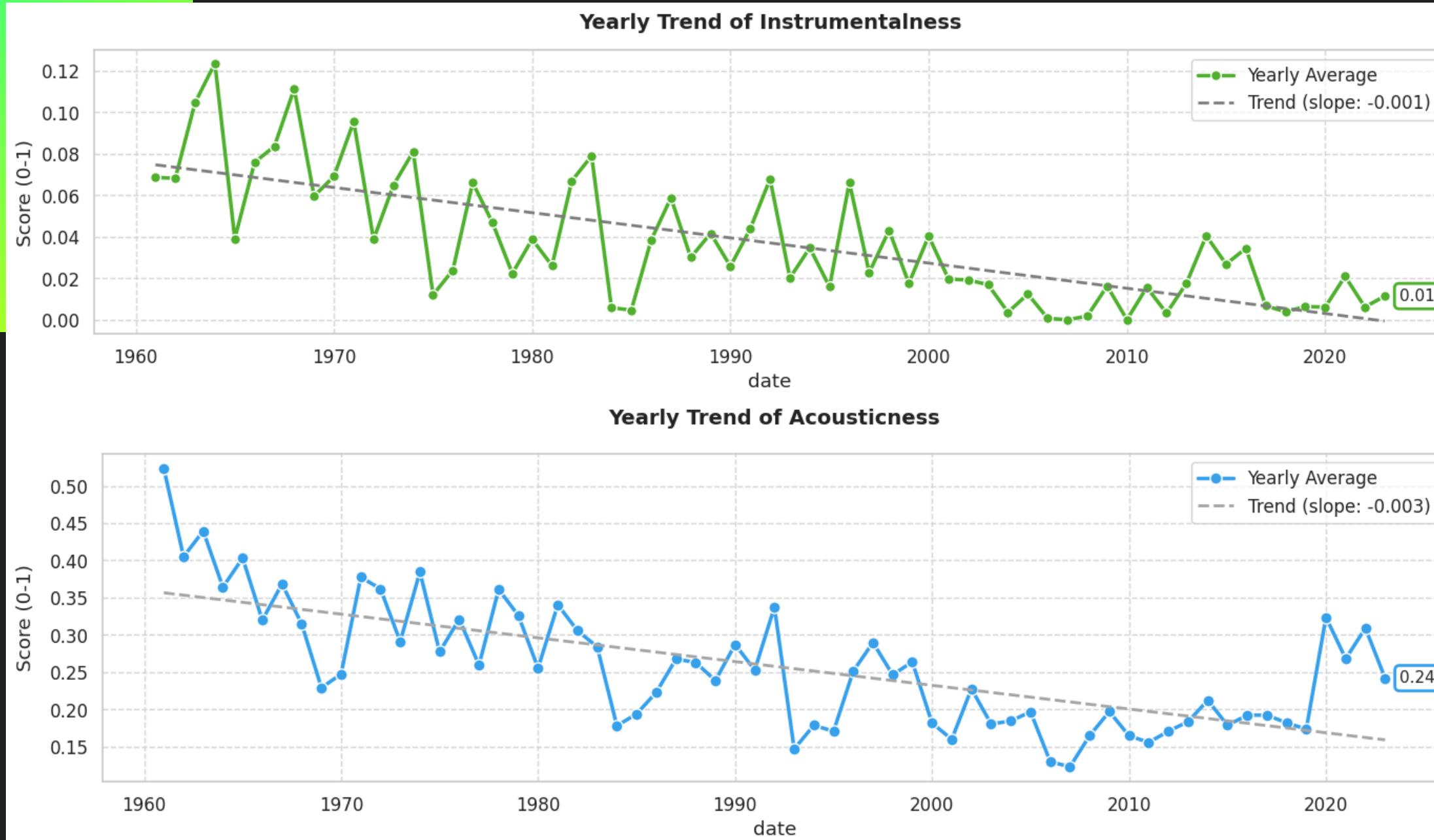
02

Mainstream music has become increasingly optimized for high-energy environments like clubs, festivals, and streaming playlists designed for engagement.

07



MUSIC TRENDS OVER TIME (PT.2)



01

A gradual decrease in instrumentalness and acousticness from 1960 to 2020

02

The rise of electronic and digital music, the decline of acoustic instruments, and the influence of studio production and streaming have reshaped popular music.



CASE STUDY: TAYLOR SWIFT

Select Artist:

Taylor Swift

Top 6 Ranking Songs

Title	Peak Rank
All Too Well (Taylor's Version)	1
Willow	1
Cardigan	1
We Are Never Ever Getting Back Together	1
Shake It Off	1
Blank Space	1

Best Rank:

1

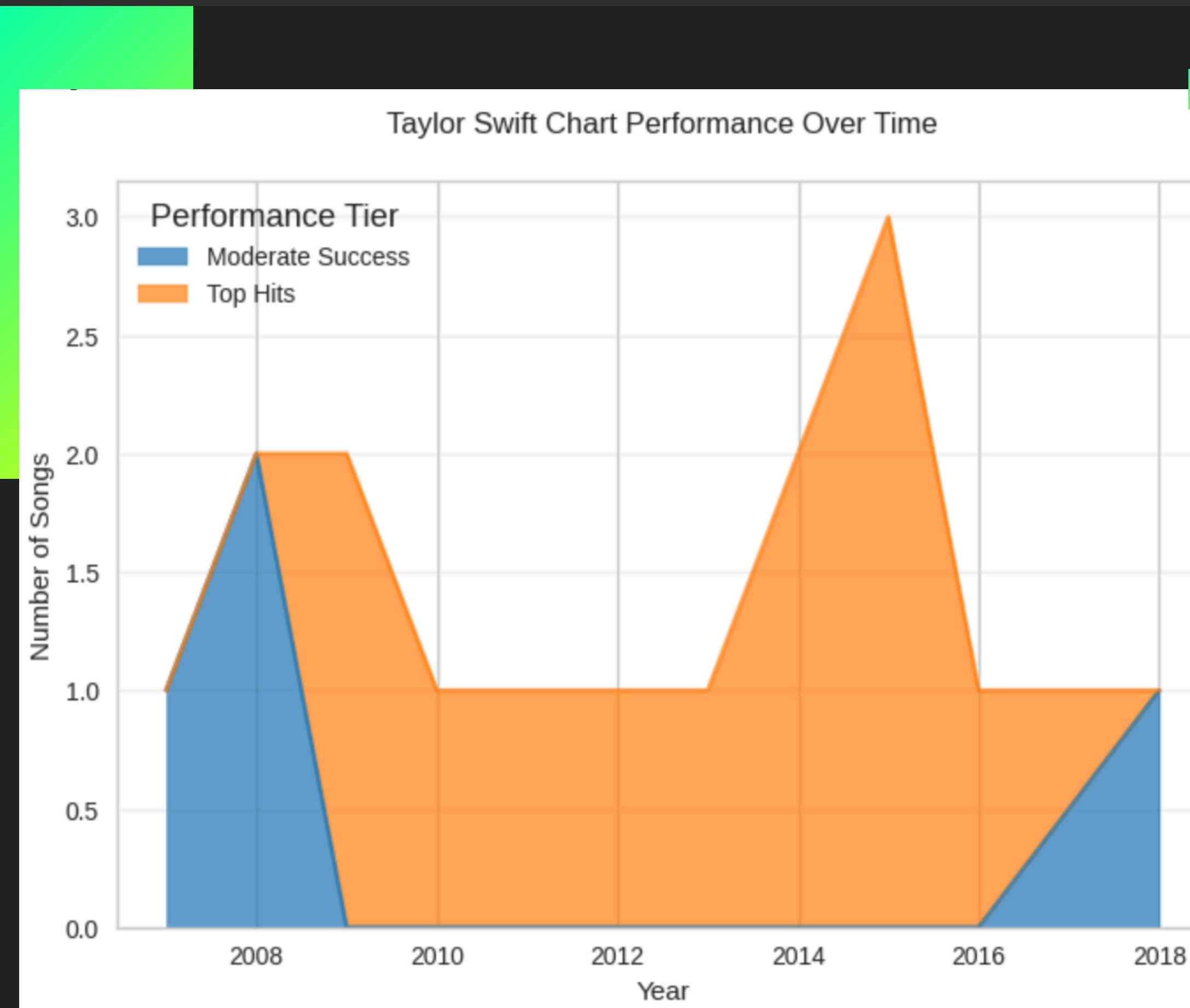
Songs on Board:

166

Longest Lasting Week:

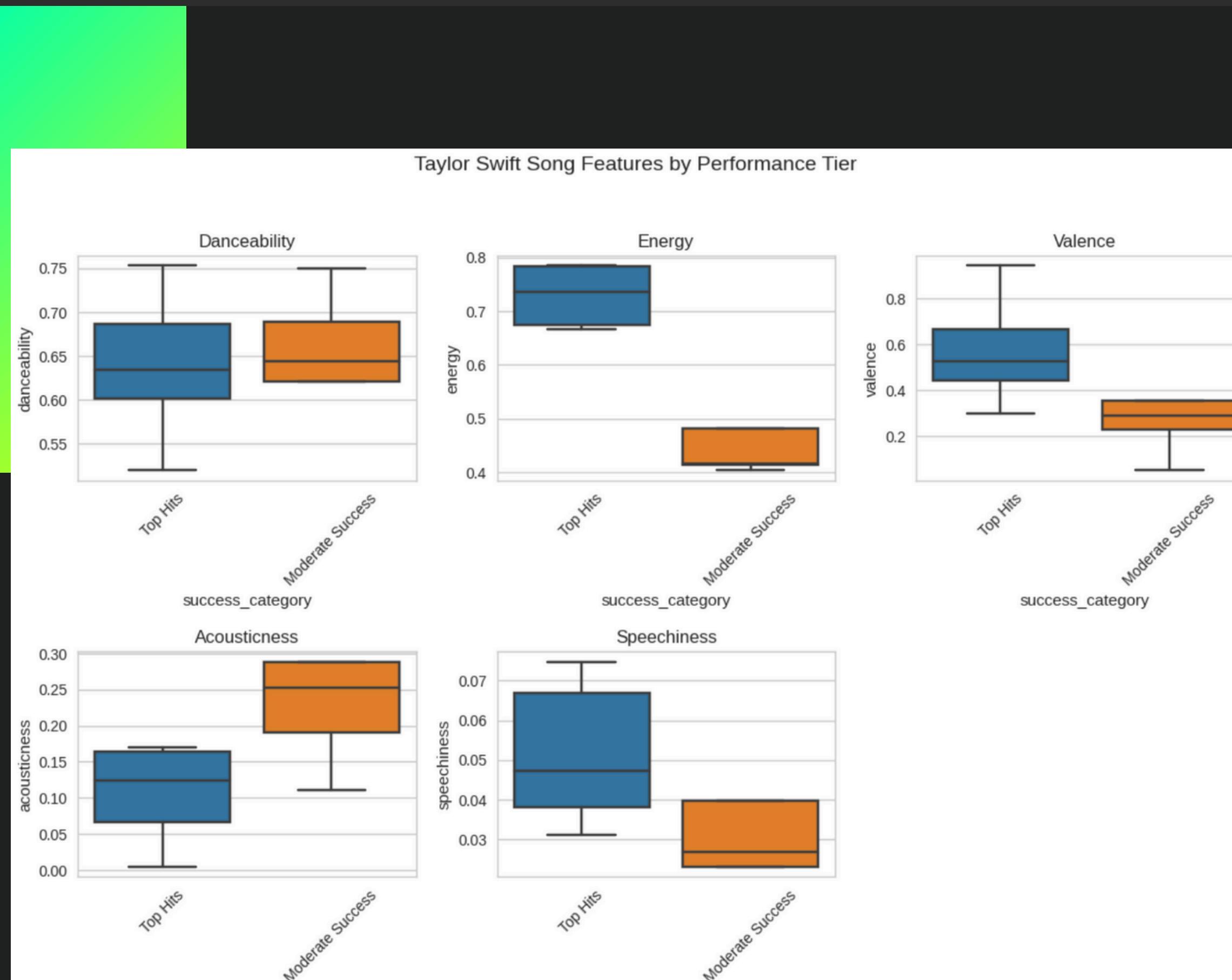
50





CASE STUDY: TAYLOR SWIFT

- Top hits: Top 10 peak position, remained on the chart for at least 10 weeks.
- Moderate success: Top 50 peak position, remained on the chart for at least 5 weeks.



CASE STUDY: TAYLOR SWIFT



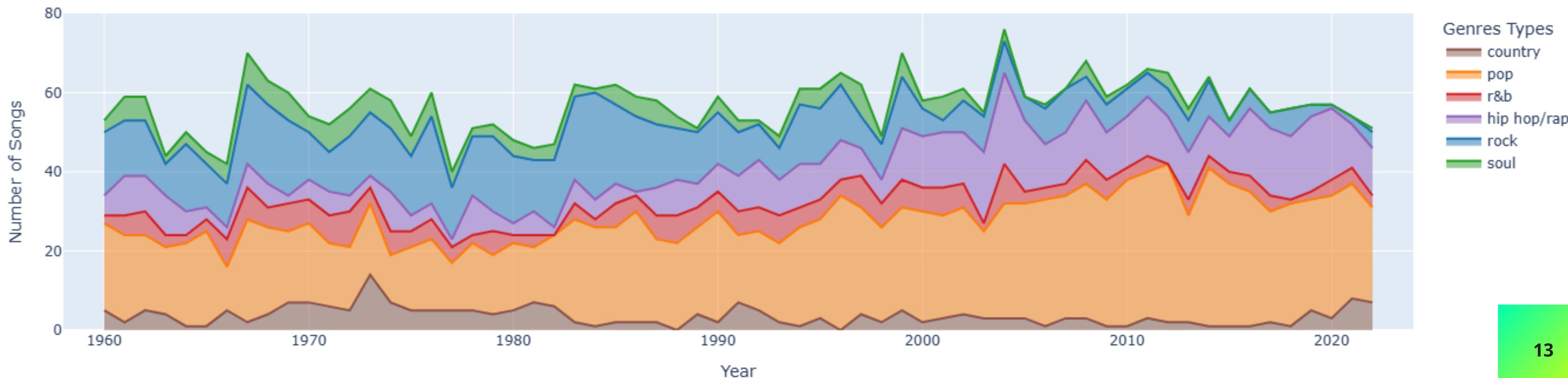
HOW HAVE MUSICAL TRENDS EVOLVED OVER TIME IN RELATION TO MAJOR INDUSTRY SHIFTS?



GENRE POPULARITY OVER TIME

- Genre Evolution Over Time: This graph illustrates the changing popularity of major music genres from 1960 to 2020, with trends showing the rise of hip hop/rap and the sustained presence of rock, pop, and R&B.
- Sub-Genre Mapping: Various sub-genres were consolidated into six main categories—country, pop, R&B, hip hop/rap, rock, and soul—to provide a clearer view of overall trends.
- Key Observations: Rock and soul peaked in the 1970s and 2000s, hip hop/rap gained dominance after the 1980s, and pop and R&B remained consistently popular over time.

Genre Popularity Over Time

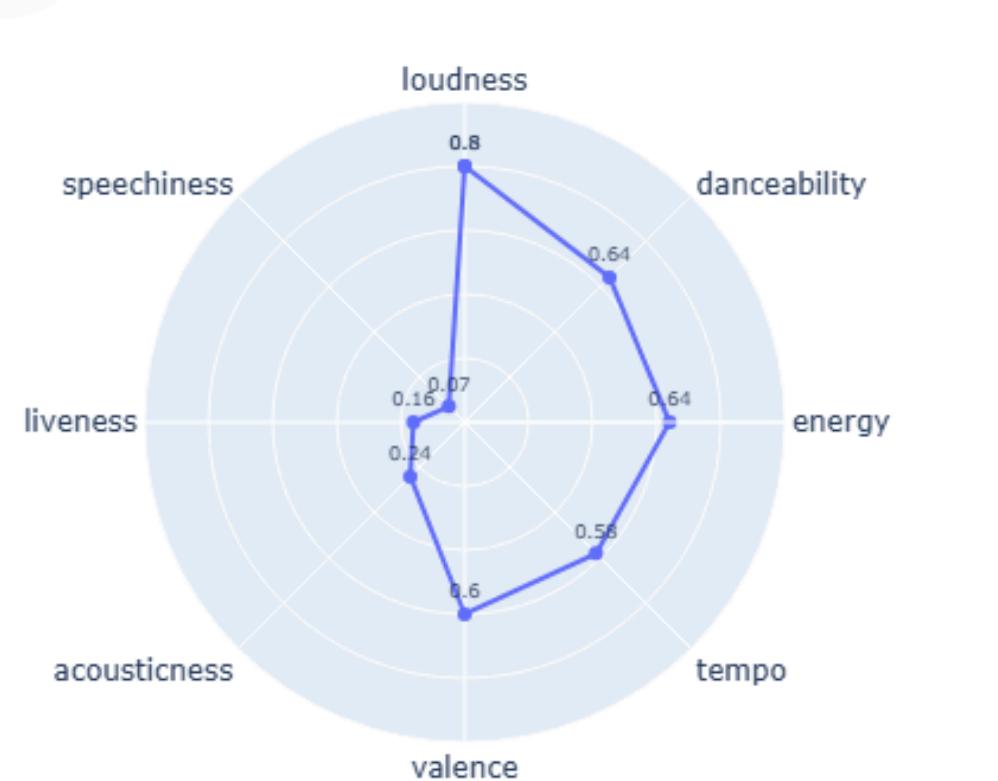




1960s

Music was more **acoustic** and had higher **liveness**, reflecting live recordings and organic instrumentation. Energy and danceability were moderate.

"Hey Jude" – The Beatles (1968)



1990s

Songs became louder and more **energetic**, with increased **danceability**, marking the rise of electronic production and pop-driven beats.

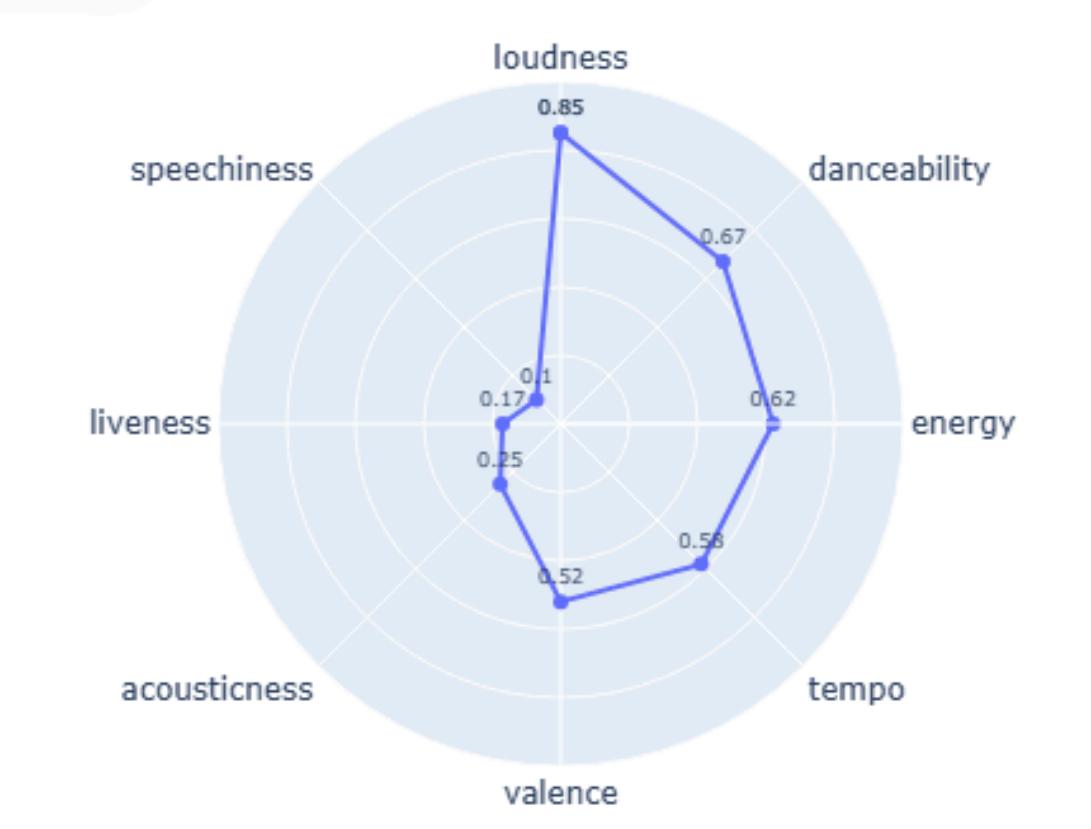
"Smells Like Teen Spirit" – Nirvana (1991)



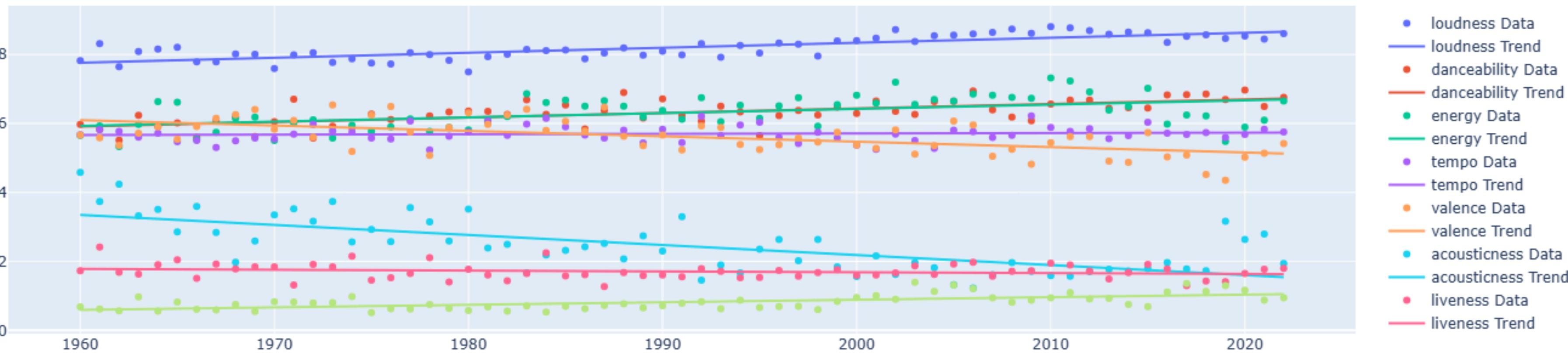
2020s

Music is highly produced, with maximum **loudness** and **danceability**, while acousticness and liveness are at their lowest, reflecting digital production trends.

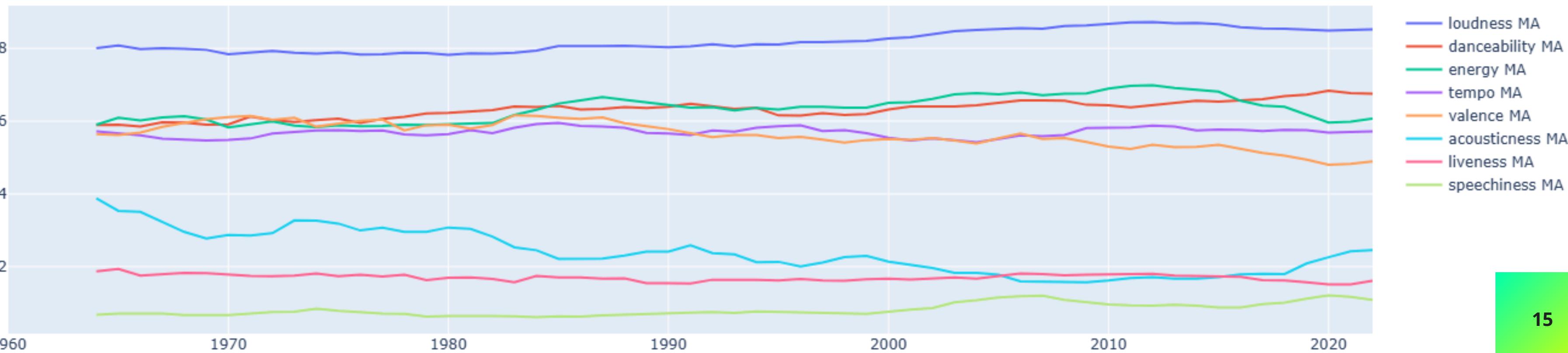
"Blinding Lights" – The Weeknd (2020)



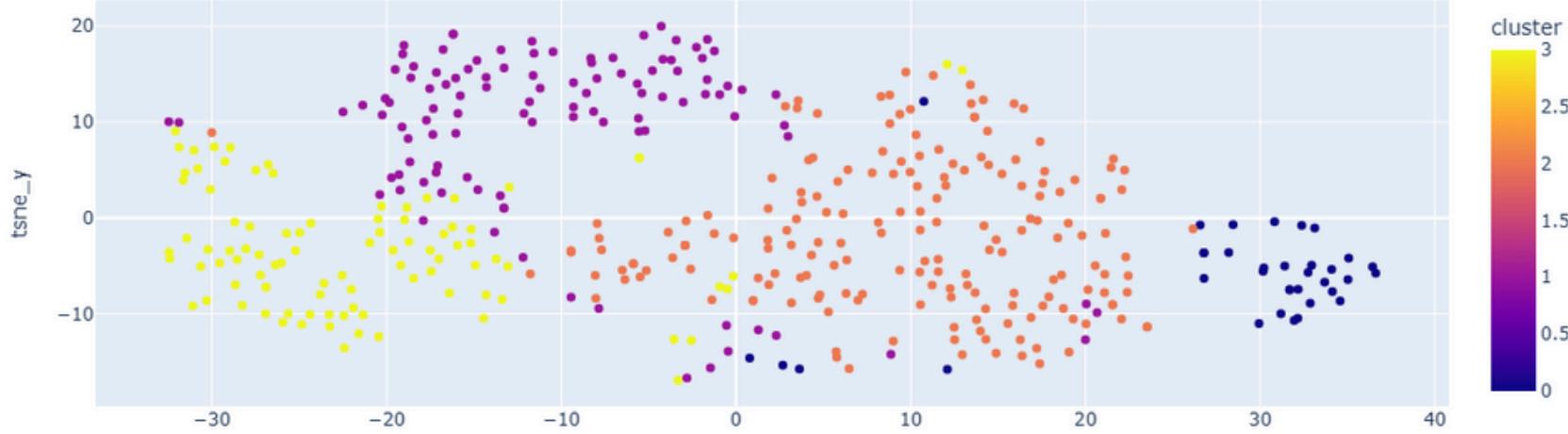
Linear Regression Trends



5-Year Moving Averages

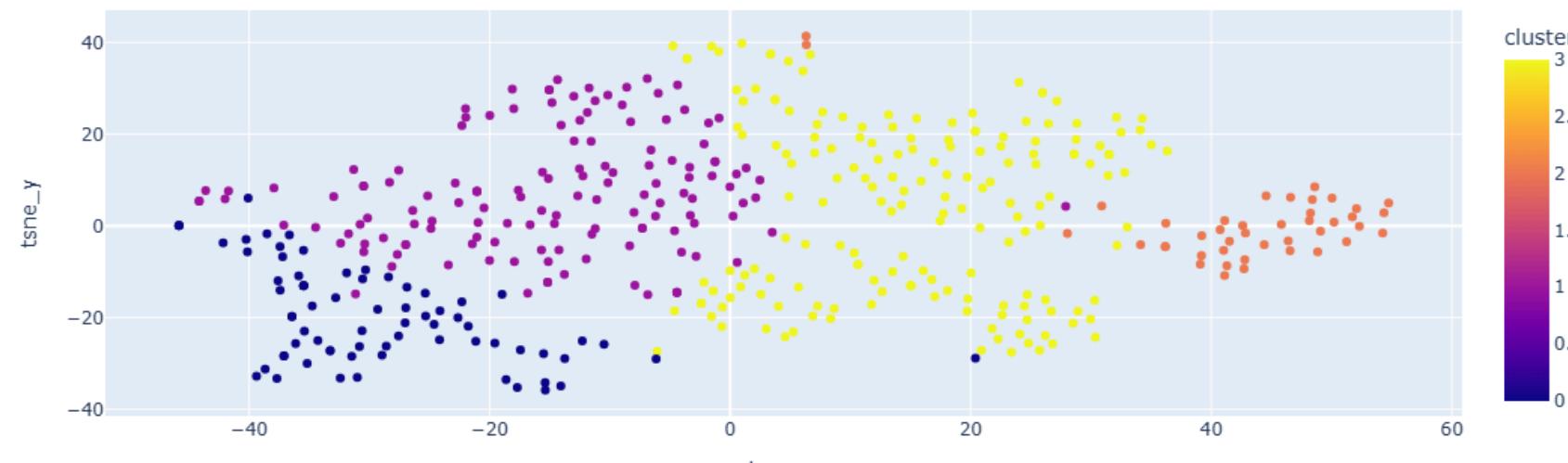


Clusters in the 1960s (t-SNE Projection)



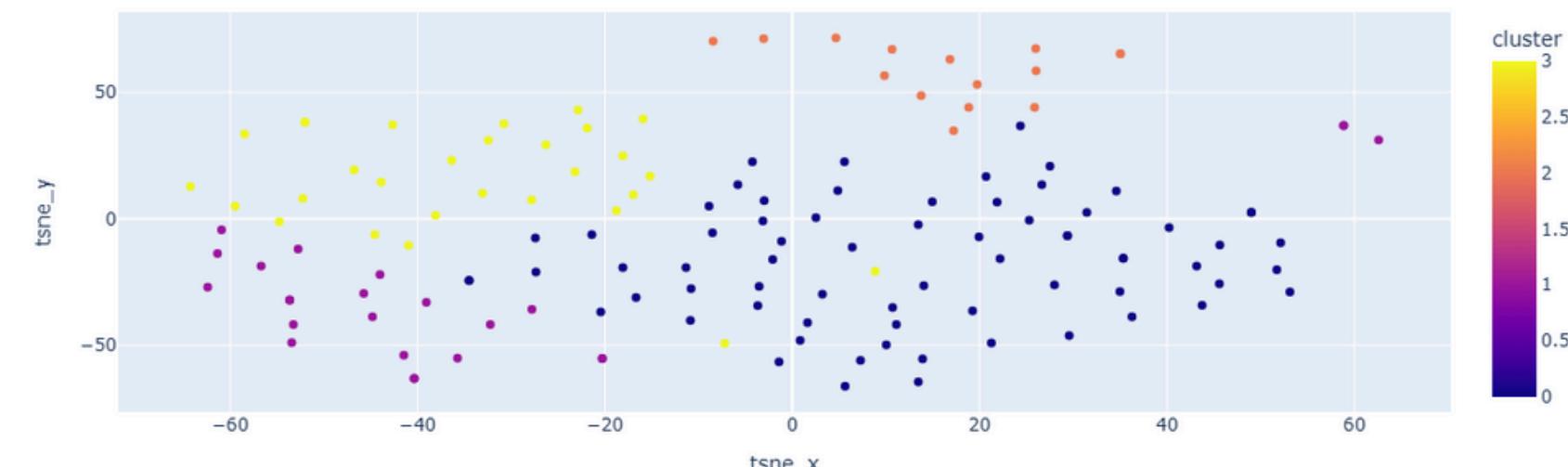
- 1960s: balanced distribution of clusters => gradual transition.
- 1990s: distinct, elongated clusters => clearer genre differentiation.
- 2020s: clusters are more fragmented => a rise in cross-genre experimentation.

Clusters in the 1990s (t-SNE Projection)



- The Rise of Latin Pop + Hip-Hop Fusion
- Post-2015: Streaming platforms made global music more accessible, and collaborations between Latin and hip-hop artists skyrocketed.

Clusters in the 2020s (t-SNE Projection)



- Examples:
 - "Despacito (Remix)" (2017)
 - "Taki Taki" (2018)
 - "Baila Baila Baila (Remix)" (2019)



THANK YOU
