

Spotify dataset analysis

An Exploratory Data Analysis [EDA] based story to find some reliable insights



Presented by
Vivek Majumdar

Data description

Column Name	Description
0 track_id	A unique identifier for the track on Spotify.
1 track_name	The title of the song.
2 artist_name	The name of the artist(s) who performed the song.
3 year	The release year of the song.
4 popularity	A measure of how popular a track is, ranging from 0 to 100.
5 artwork_url	A URL pointing to the album artwork for the track.
6 album_name	The name of the album the track belongs to.
7 acousticness	A confidence measure indicating whether the track is acoustic.
8 danceability	A measure of how suitable a track is for dancing.
9 duration_ms	The duration of the track in milliseconds.
10 energy	A perceptual measure of intensity and activity.
11 instrumentalness	Predicts whether a track contains no vocal content.
12 key	The key the track is in, represented as an integer (0-11).
13 liveness	Detects the presence of an audience in the recording.
14 loudness	The overall loudness of a track in decibels (dB).
15 mode	Indicates the modality (major or minor) of a track.
16 speechiness	A measure detecting the presence of spoken words.
17 tempo	The overall estimated tempo of a track in beats per minute (BPM).
18 time_signature	An estimated overall time signature of a track.
19 valence	A measure from -1.0 to 1.0 describing the musical positiveness contained in a track.
20 track_url	A URL to the Spotify track.
21 language	The detected language of the song's lyrics.

Popularity

A measure of how popular a track is, ranging from 0 to 100

Acousticness

A confidence measure indicating whether the track is acoustic, ranging from -1.0 to 1.0

Energy

A perceptual measure of intensity and activity, ranging from -1.0 to 1.0

Tempo

The overall estimated tempo of a track in beats per minute (BPM)

Danceability

A measure of how suitable a track is for dancing, ranging from -1.0 to +1.0

Duration

The duration of the track (here, in milliseconds)

Univariate Analysis

01

Popularity

A measure of how popular a track is, ranging from 0 to 100

02

Tempo

The overall estimated tempo of a track in beats per minute (BPM)

03

Duration

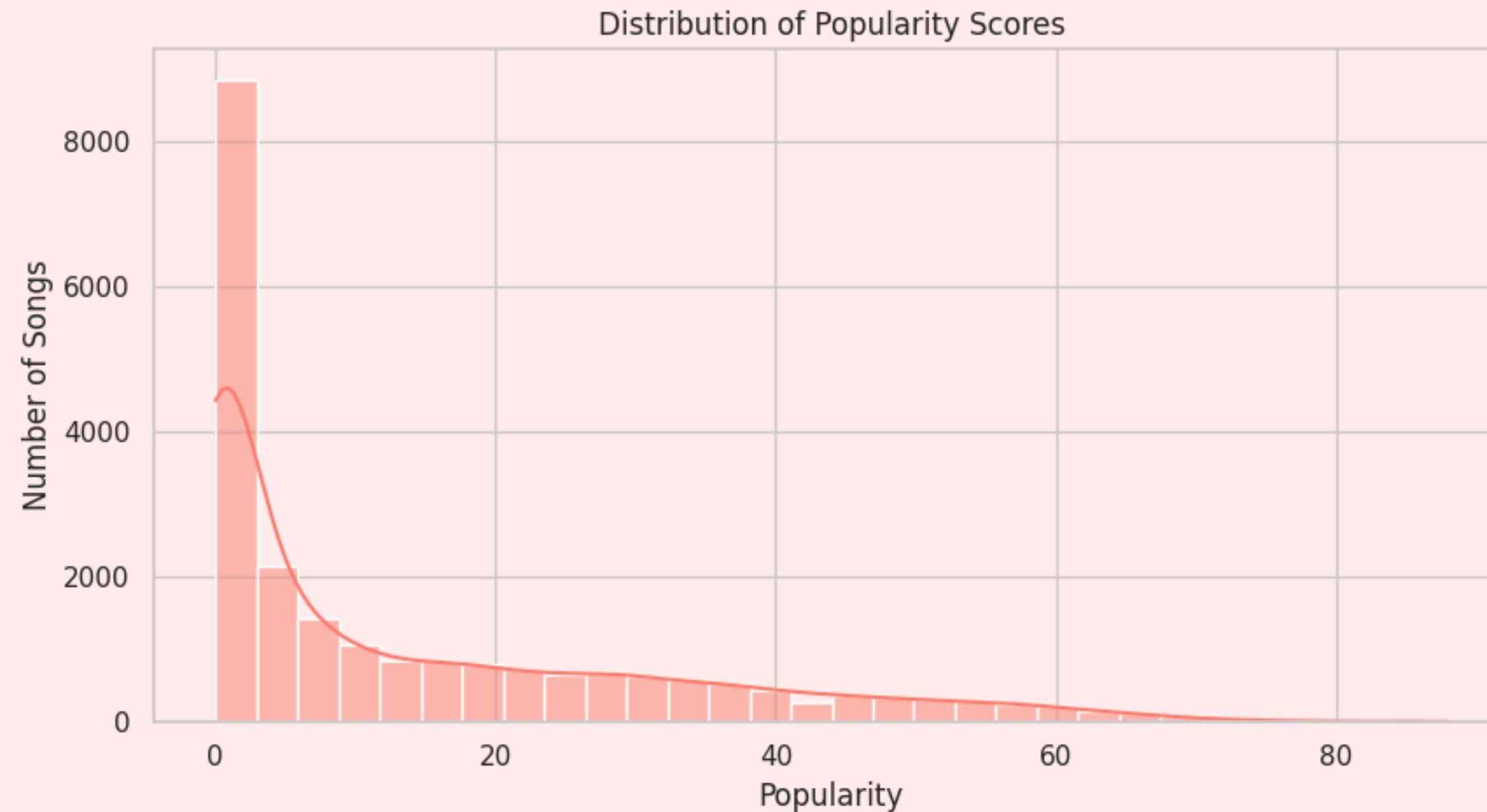
The duration of the track (here, in milliseconds)

04

Danceability

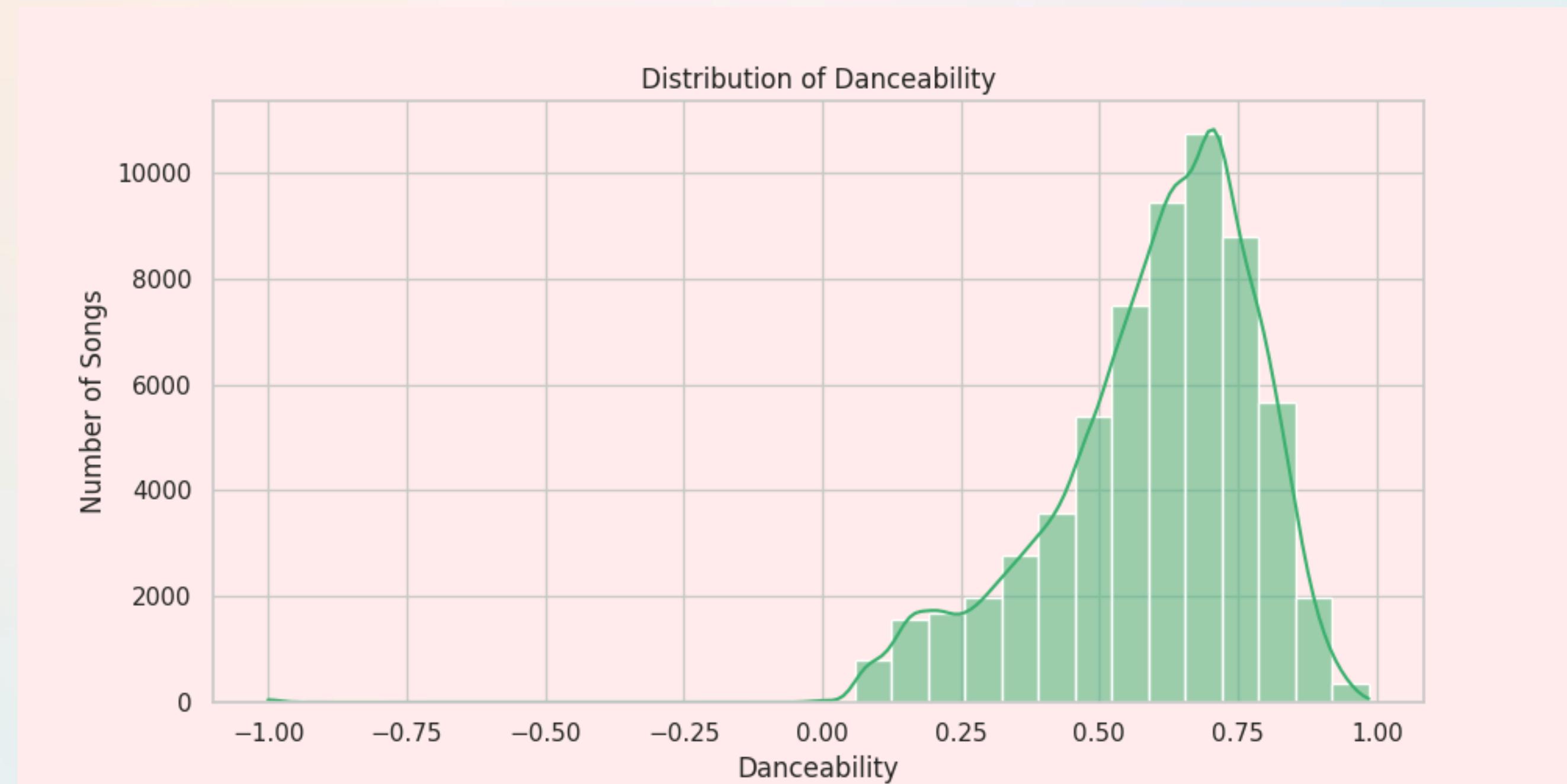
A measure of how suitable a track is for dancing, ranging from -1.0 to +1.0

Popularity scores



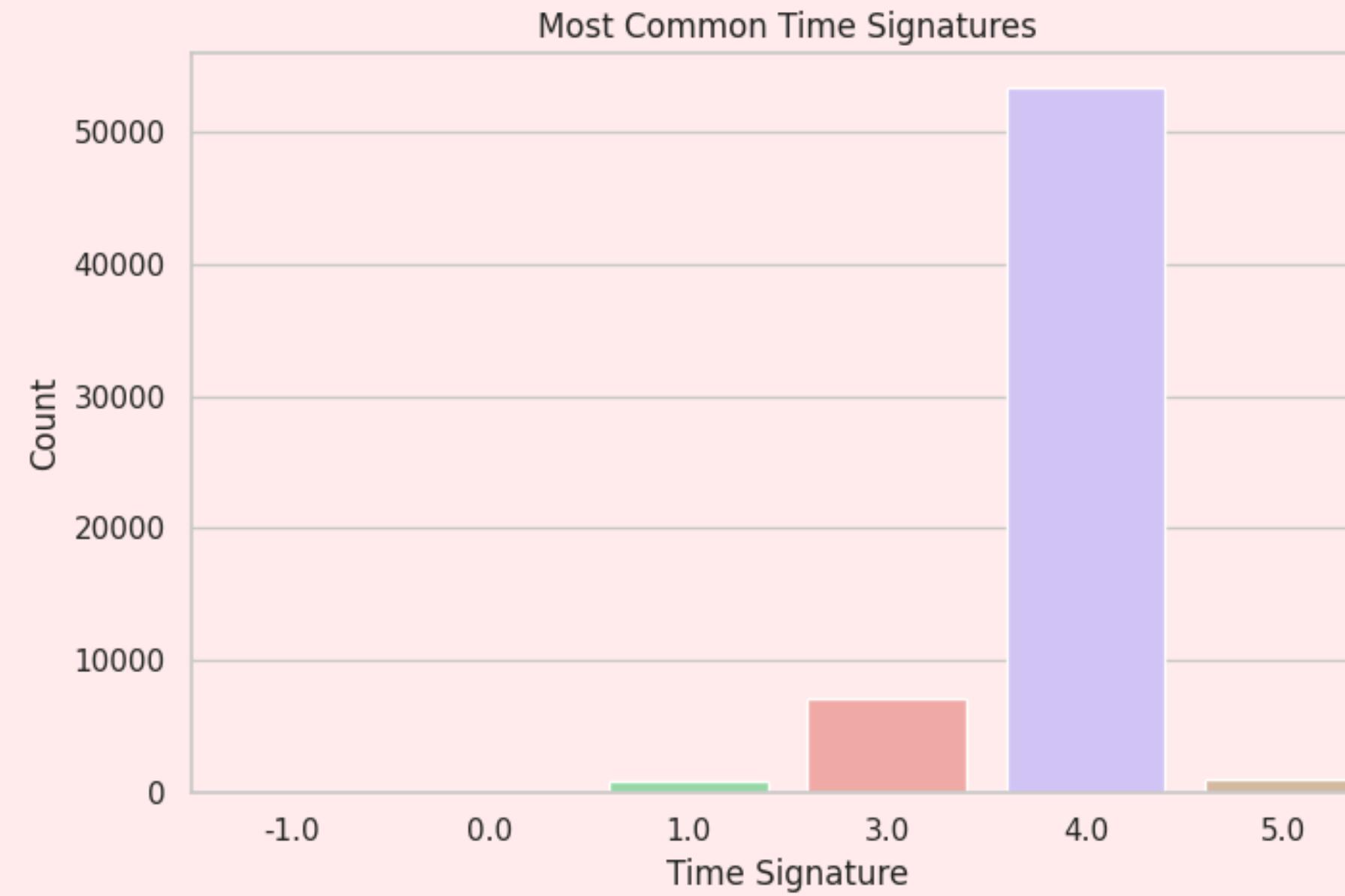
Most songs have low to moderate popularity (0-50). Very few tracks are highly popular.

Danceability distribution



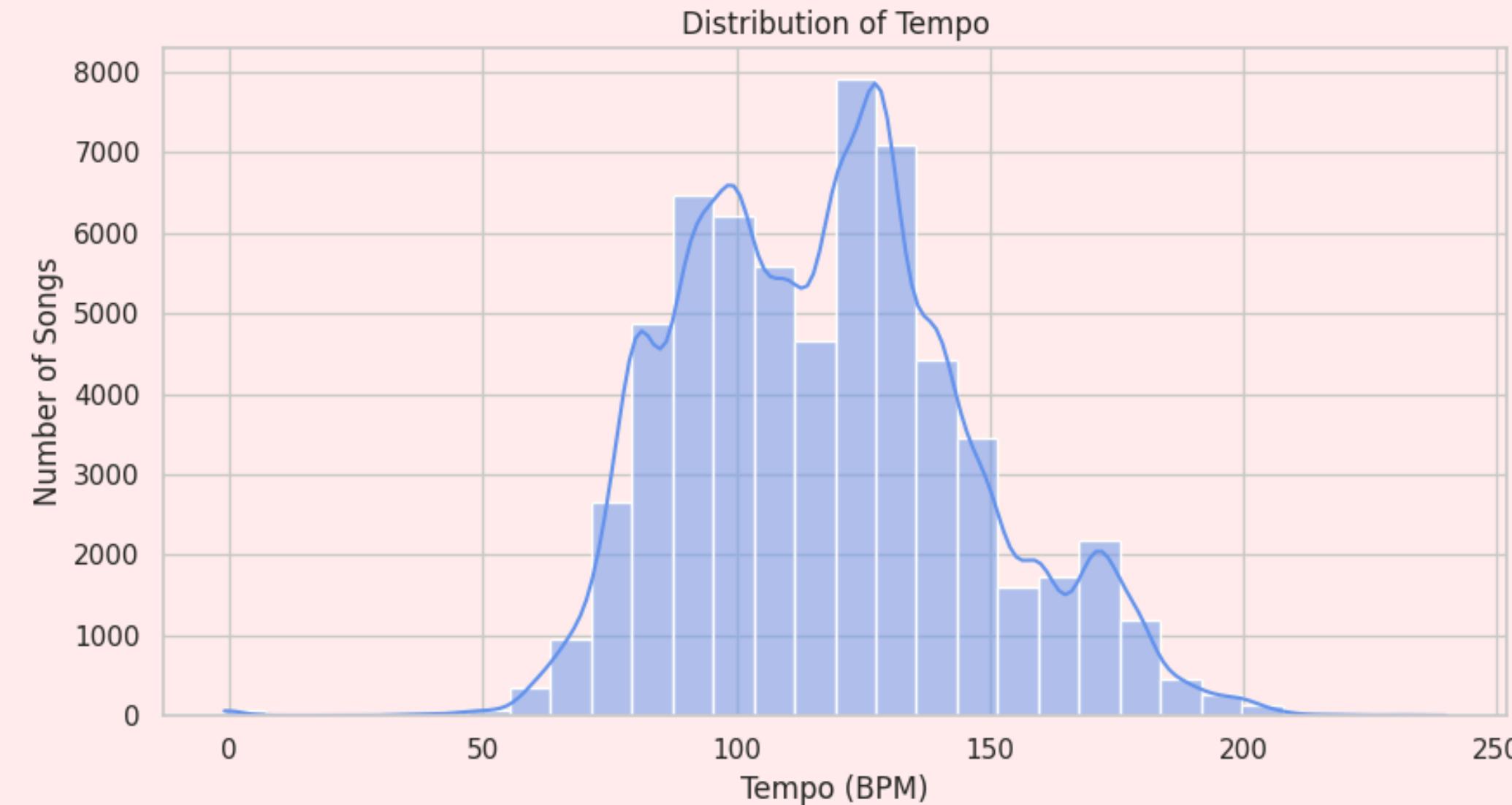
Danceability is mostly moderate to high (0.5-0.8). Aiming for a balanced danceability (0.6-0.8) might align with listener preferences.

Most common time signatures



4/4 is the dominant time signature. 4/4 is the safest and most accepted. Experimental time signatures can still work, but may cater to niche audiences

Distribution of Tempo



Most tempos fall between 90-140 BPM, peaking around 120 BPM. For commercial success, especially in pop or dance genres, targeting tempos in the 100-130 BPM range is a good starting point. It aligns with natural human rhythm and is dancefloor-friendly

Insights and Implications:

- Most songs have low to moderate popularity (0-50). Very few tracks are highly popular.
- Danceability is mostly moderate to high (0.5-0.8). Aiming for a balanced danceability (0.6-0.8) might align with listener preferences.
- 4/4 is the dominant time signature. 4/4 is the safest and most accepted. Experimental time signatures can still work, but may cater to niche audiences.
- Most tempos fall between 90-140 BPM, peaking around 120 BPM. For commercial success, especially in pop or dance genres, targeting tempos in the 100-130 BPM range is a good starting point. It aligns with natural human rhythm and is dancefloor-friendly

Bivariate analysis

01

Duration vs Popularity

Clearly identify the focus area or project for the SWOT analysis.

02

Popularity-Danceability relation

Collect relevant information about internal strengths/weaknesses and external opportunities/threats.

03

Popularity-Energy relation

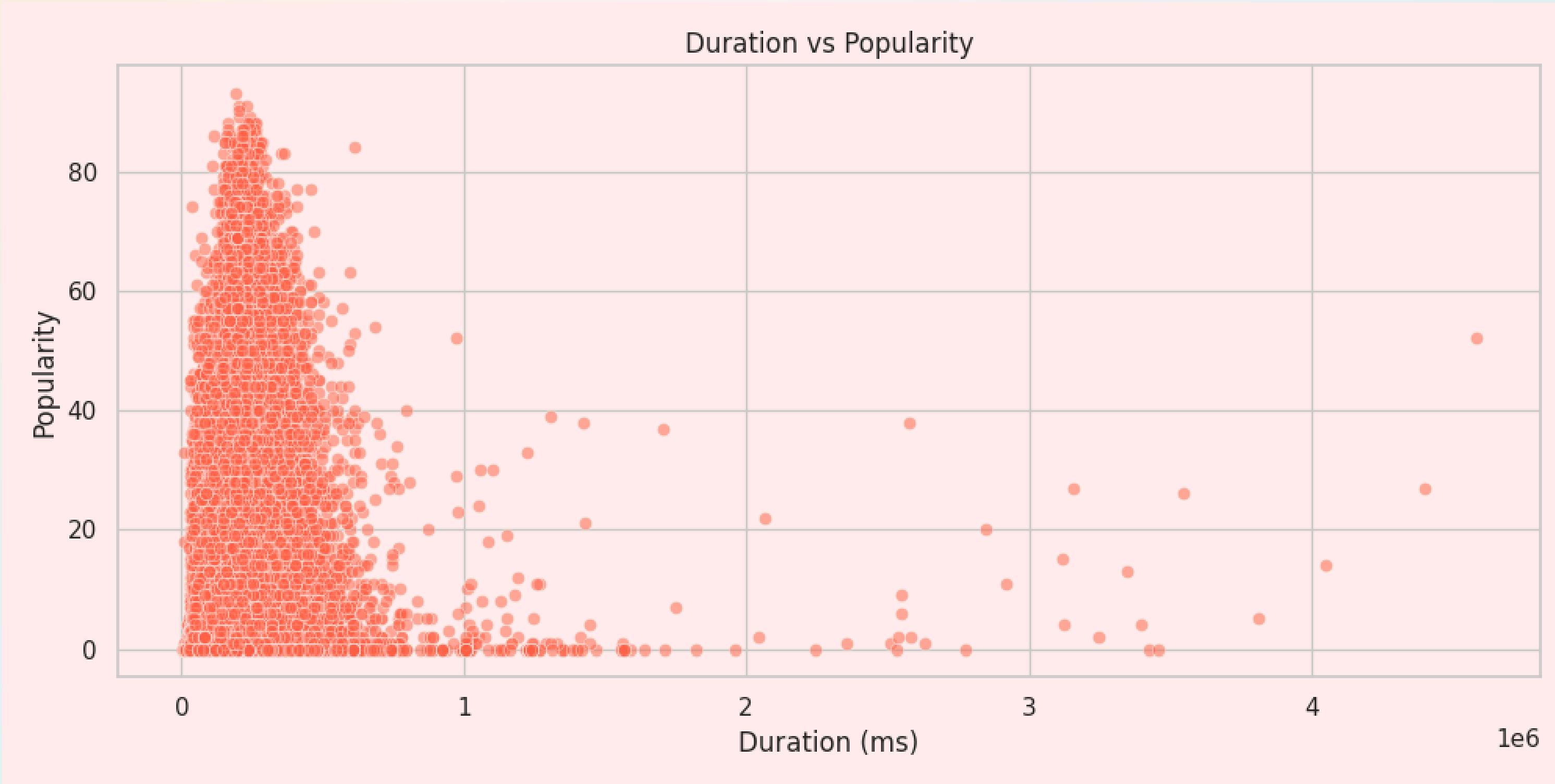
Thoroughly review the data to uncover meaningful patterns and strategic implications.

04

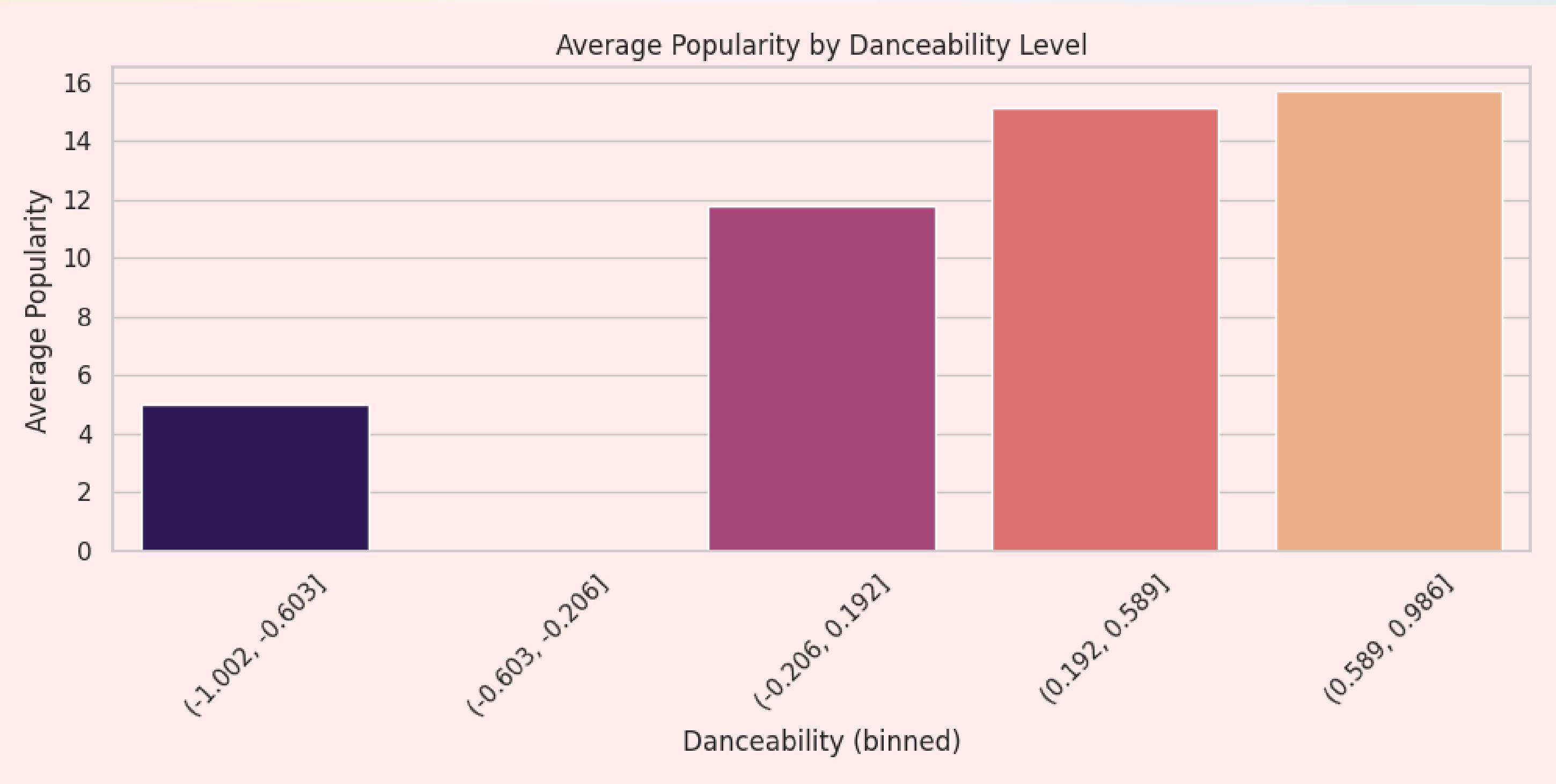
Popularity-Acousticness relation

Determine the most critical strengths, weaknesses, opportunities, and threats to address.

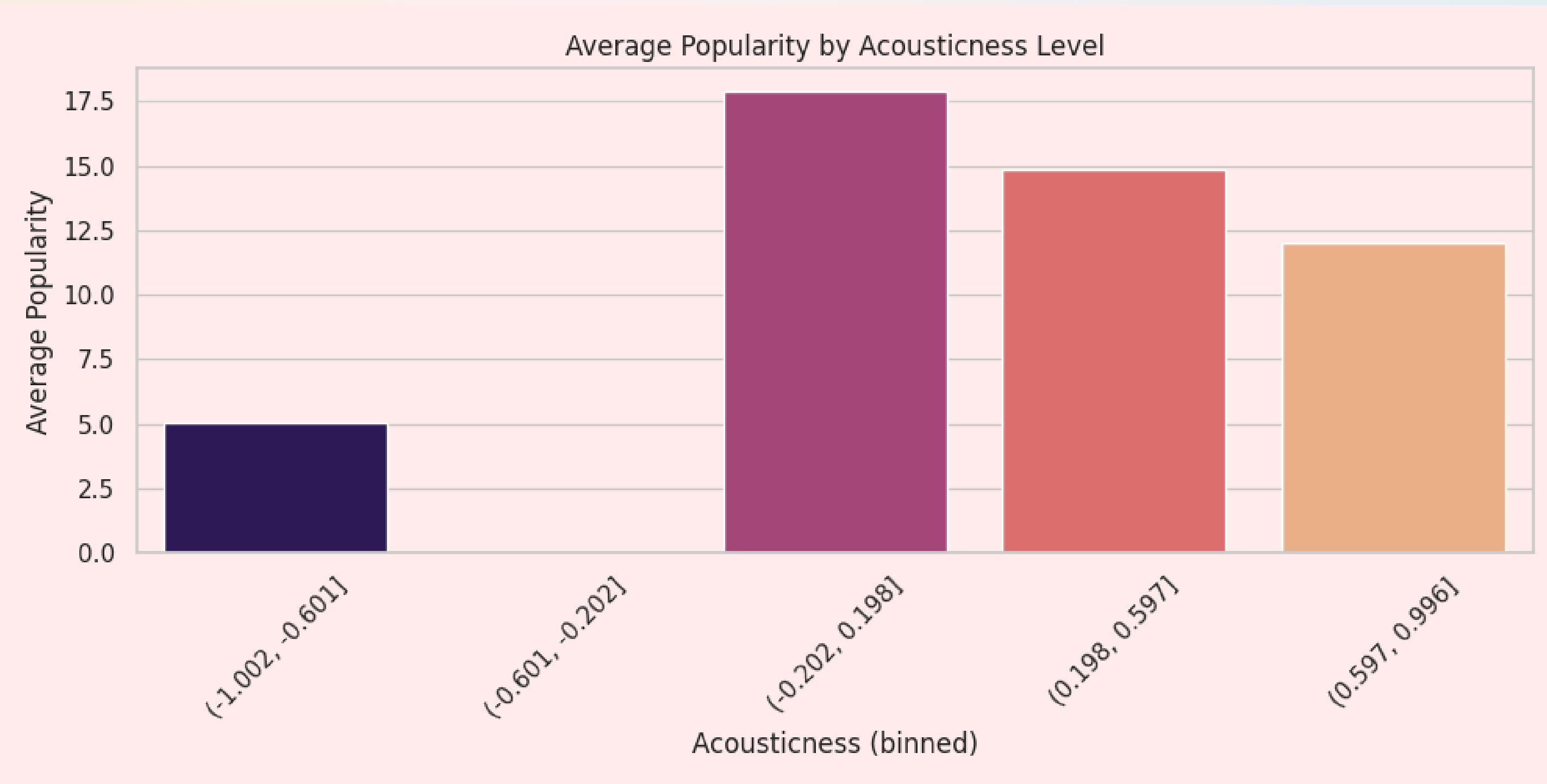
Duration vs Popularity



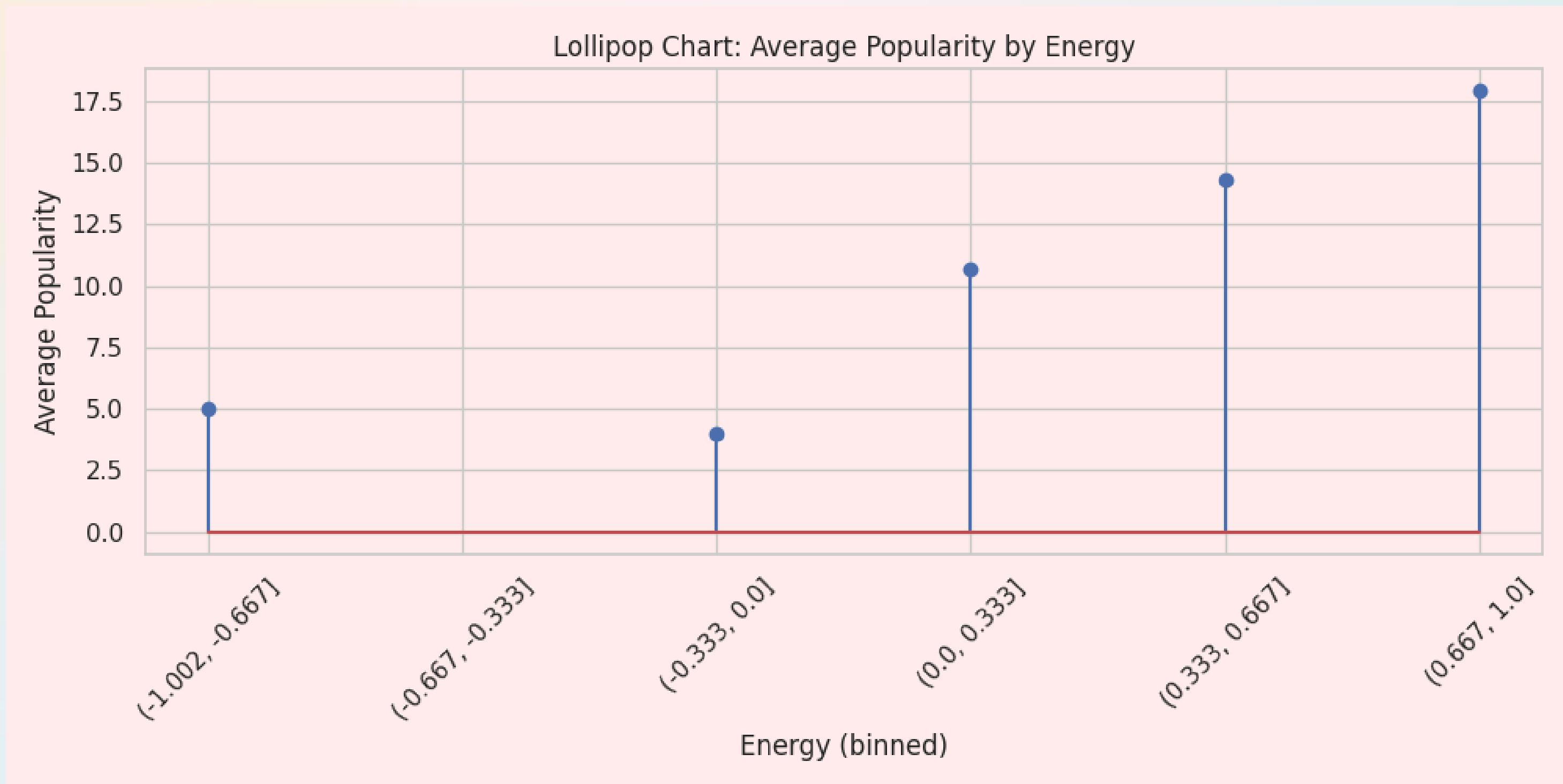
Popularity-Danceability relation



Popularity-Acousticness relation



Popularity-Energy relation



Insights and Implications:-

- There is no strong positive or negative trend. Tracks with both high and low acousticness can be popular, but very high acousticness (above 0.8) is less common among top hits.
- popular tracks tend to have high danceability and energy, and moderate valence (emotional positivity). Energetic, danceable tracks that evoke positive emotions are more likely to be popular.
- Song releases peaked around 2020, then declined. Implication (as a producer): Competition is fierce around 2020-2021; releasing now might help cut through less noise if quality stands out.
- Popularity has stayed steady over time. Implication: Timeless sound matters—focus on strong songwriting and hooks over trends.
- Danceability is trending upward. Implication: Prioritize rhythm and groove; tracks with strong bounce/connectivity may gain better traction.

Time Series Analysis

01

Number of Tracks - Year

Is there a increase or decrease in no. of songs released ?

02

Average Popularity - Year

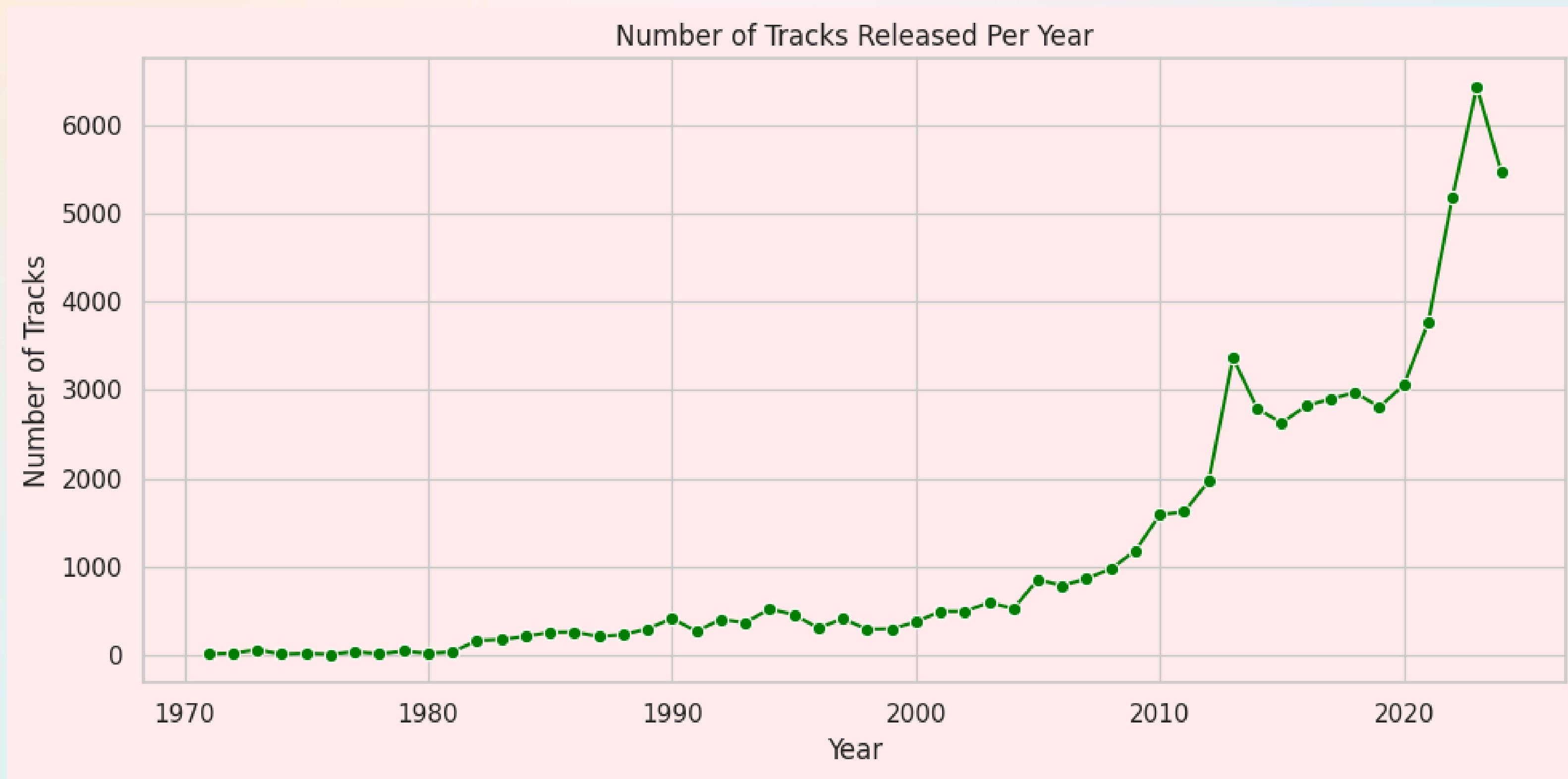
Do the songs remain popular with time ?

03

Danceability-Energy Trend

Is there a change in the trend over time ?

Track Released per Year

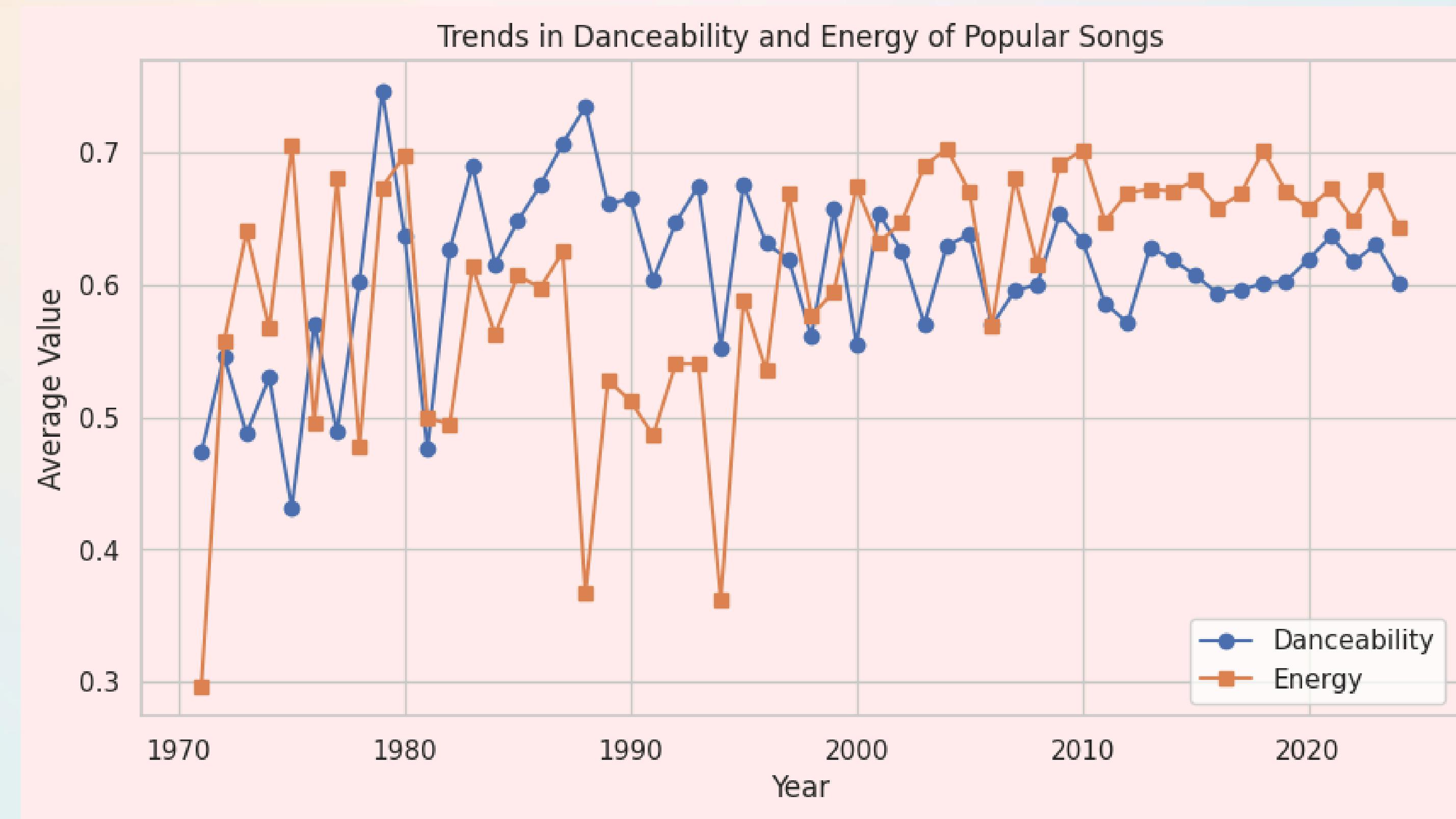


Average Popularity per Year

Average Popularity of Songs Over the Years



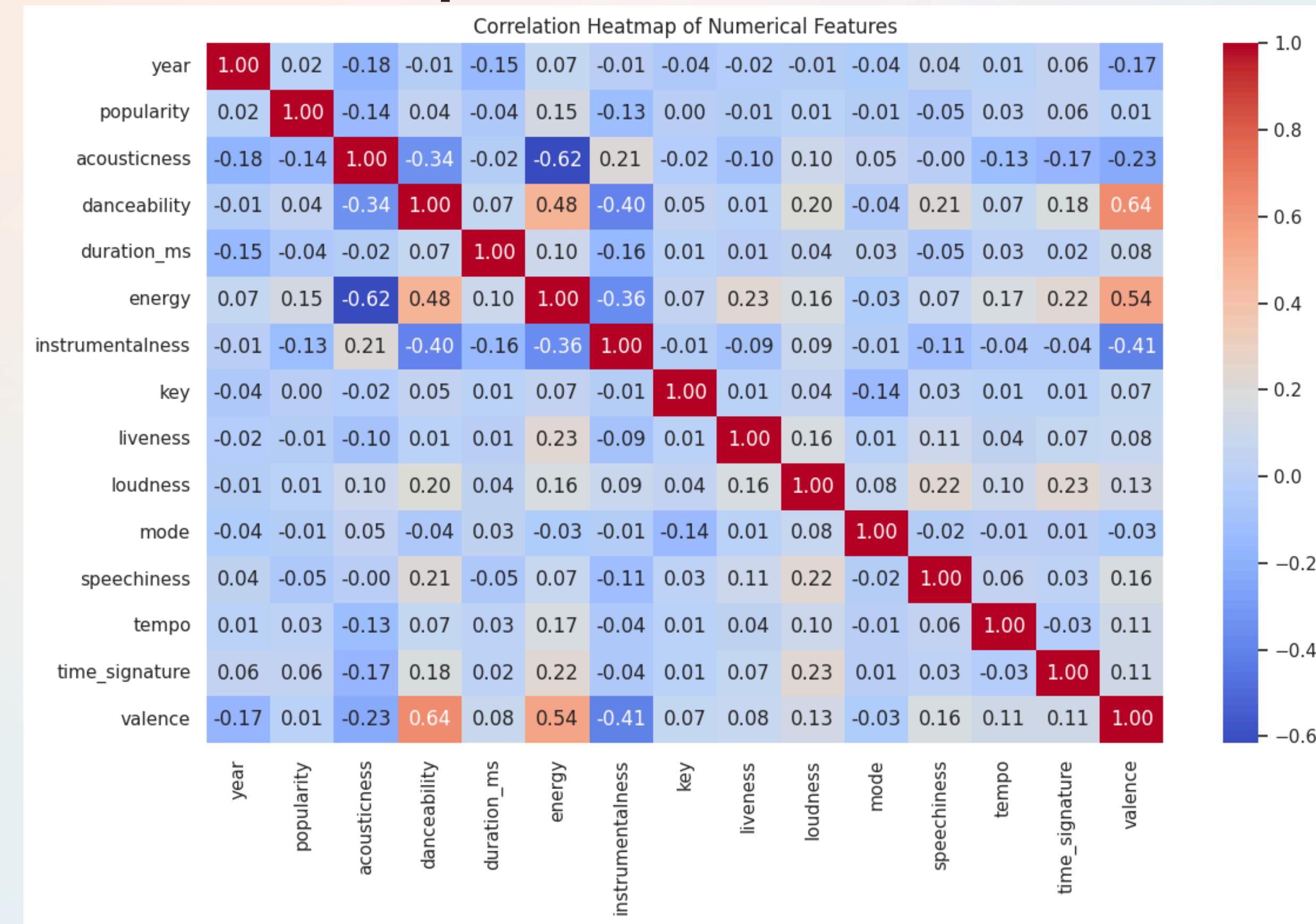
Trends in Danceability and Energy



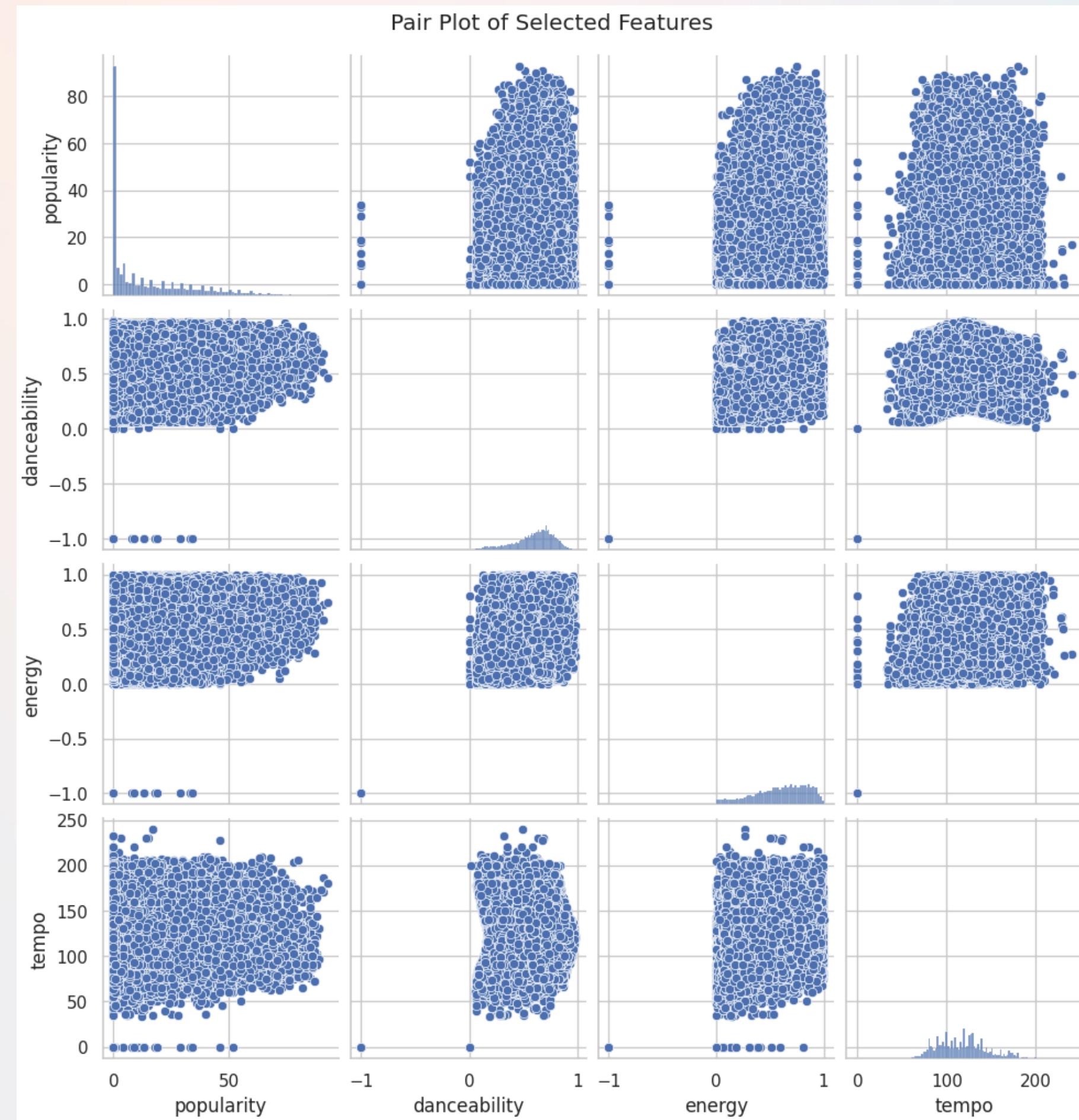
Insights and Implications:-

- Song releases peaked around 2020, then declined. Competition is fierce around 2020-2021; releasing now might help cut through less noise if quality stands out.
- Insight: Song releases peaked around 2020, then declined. Implication (as a producer): Competition is fierce around 2020-2021; releasing now might help cut through less noise if quality stands out
- Danceability is trending upward. Implication: Prioritize rhythm and groove; tracks with strong bounce/connectivity may gain better traction.
- Energy levels are relatively constant. Implication: Maintain consistent intensity—high energy tracks still work, but don't overproduce at the cost of clarity.
- Acousticness is falling. Implication: Modern tracks lean electronic—use synths, FX, and digital layers, but keep acoustic elements for emotional depth if needed.

Correlation Heatmap



Pair Plot of Features



Insights and Implications

- Loud, energetic, and positive songs tend to go together. Implication: Boost loudness and energy for high-impact, feel-good tracks.
- Popular songs cluster around moderate-to-high energy and danceability. Implication: Balance groove and energy; tempo matters less than feel.

Thank You!

