SPOTIFY TRACKS ANALYSIS

To explore and analyze the Spotify dataset to uncover patterns and trends that define popular music through detailed Exploratory Data Analysis (EDA)

PRESENTED BY
NAYANIKA CHATTERJEE
COMPUTER SCIENCE & ENGINEERING
3RD SEM

INTRODUCTION

The Spotify dataset provides detailed information about tracks, including their audio features such as danceability, energy, tempo, acousticness, instrumentalness, speechiness, loudness, valence, and popularity. This project performs an Exploratory Data Analysis (EDA) to uncover trends in music production and popularity.
The objective is to identify what makes a track popular, how audio features interact, and how popular music has evolved over time.

METHODOLOGY

Dataset Used: Cleaned Spotify dataset (containing features and popularity scores).

Approach:

- Univariate Analysis: To understand individual feature distributions.
- Bivariate Analysis: To study relationships between two features and popularity.
- Multivariate Analysis: To identify combinations of features influencing success.
- Time Series Analysis: To track trends in features and popularity across years.

Tools & Libraries: Python, Pandas, Matplotlib, Seaborn, Pygam, Statsmodels, Scipy, Sklearn.

STEPS FOLLOWED BEFORE GRAPHICAL ANALYSIS:-

1. Importing Libraries
2. Loading and Reading Data
3. Basic Data Overview
4. Checking and Handling Missing Values
5. Removing Duplicate Records
6. Outlier Detection (IQR Method)
7. Fixing Datatypes
8. Saving Cleaned Data

SPOTIFY DATA DESCRIPTION

- **track_id**: A unique identifier for the track on Spotify
- **track_name**: The title of the song
- **artist_name**: The name of the artist(s) who performed the song
- **year**: The release year of the song
- **popularity**: A measure of how popular a track is, ranging from 0 to 100
- **artwork_url**: A URL pointing to the album artwork for the track
- **album_name**: The name of the album the track belongs to
- **acousticness**: A confidence measure indicating whether the track is acoustic, ranging from -1.0 to 1.0
- **danceability**: A measure of how suitable a track is for dancing, ranging from -1.0 to +1.0
- **duration_ms**: The duration of the track in milliseconds
- **energy**: A perceptual measure of intensity and activity, ranging from -1.0 to 1.0
- **instrumentalness:** Predicts whether a track contains no vocal content, ranging from -1.0 to +1.0
- **key**: The key the track is in, represented as an integer (e.g., 0 = C, 1 = C#, etc.)
- **liveness**: Detects the presence of an audience in the recording, ranging from -1.0 to +1.0
- **loudness**: The overall loudness of a track in decibels (dB)
- **mode**: Indicates the modality (major or minor) of a track (0 for minor, 1 for major)
- **speechiness**: A measure detecting the presence of spoken words in a track
- **tempo**: The overall estimated tempo of a track in beats per minute (BPM)
- **time_signature**: An estimated overall time signature of a track
- **valence**: A measure from -1.0 to 1.0 describing the musical positiveness conveyed by a track
- **track_url**: A URL to the Spotify track
- **language**: The detected language of the song's lyrics

UNIVARIATE ANALYSIS

Univariate analysis involves examining one feature or variable at a time to understand its distribution and key characteristics. This analysis provides insights into the overall structure of the dataset by identifying trends such as the most common musical keys, the typical tempo range, or the general levels of danceability and energy. It helps establish a foundational understanding of how individual features behave before exploring their relationships with others.

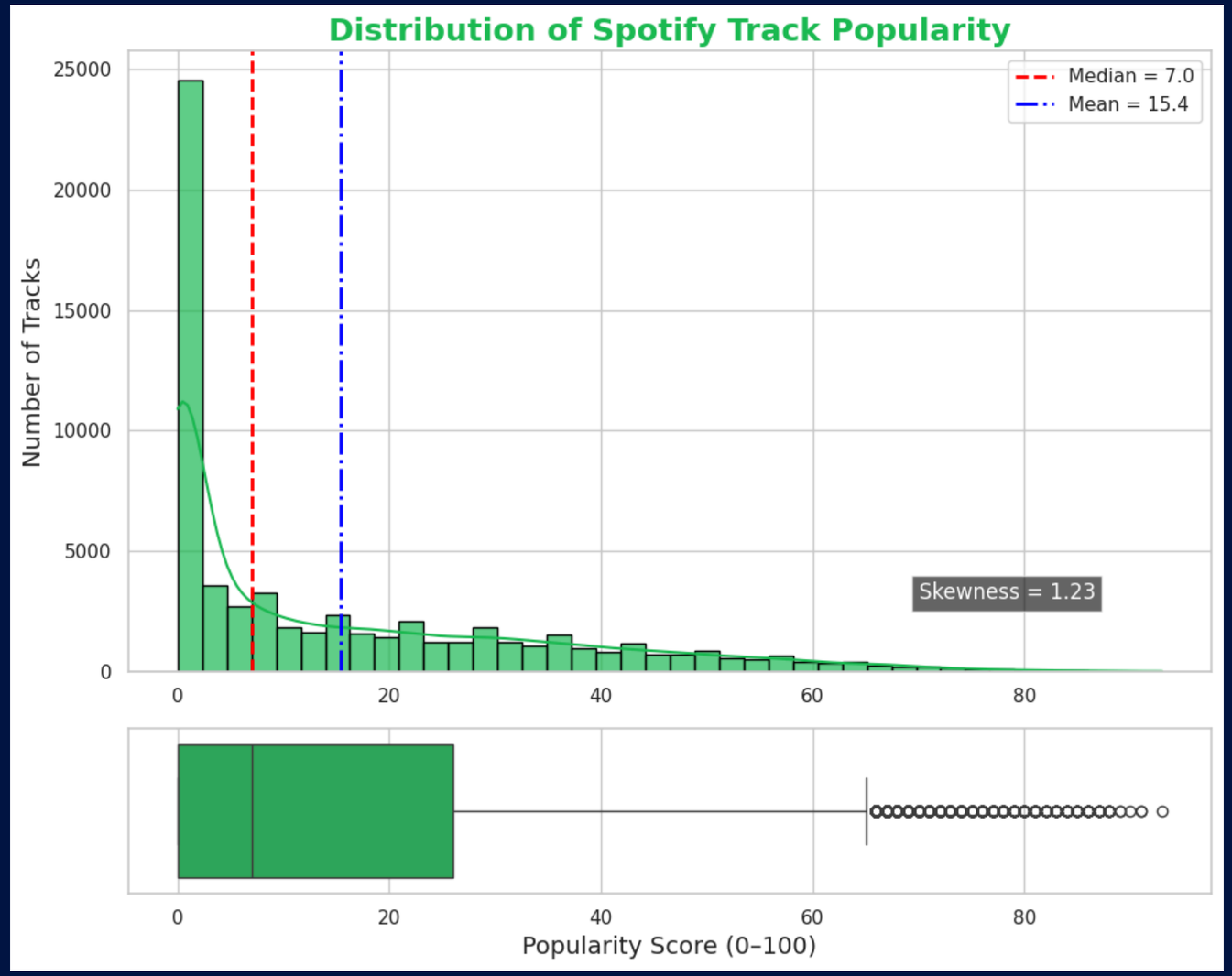
1. WHAT IS THE OVERALL DISTRIBUTION OF POPULARITY SCORES ACROSS ALL TRACKS IN THE DATASET? (ARE MOST SONGS MODERATELY POPULAR, OR IS IT SKEWED TOWARDS VERY HIGH/LOW POPULARITY?)

Breakdown of track popularity categories:

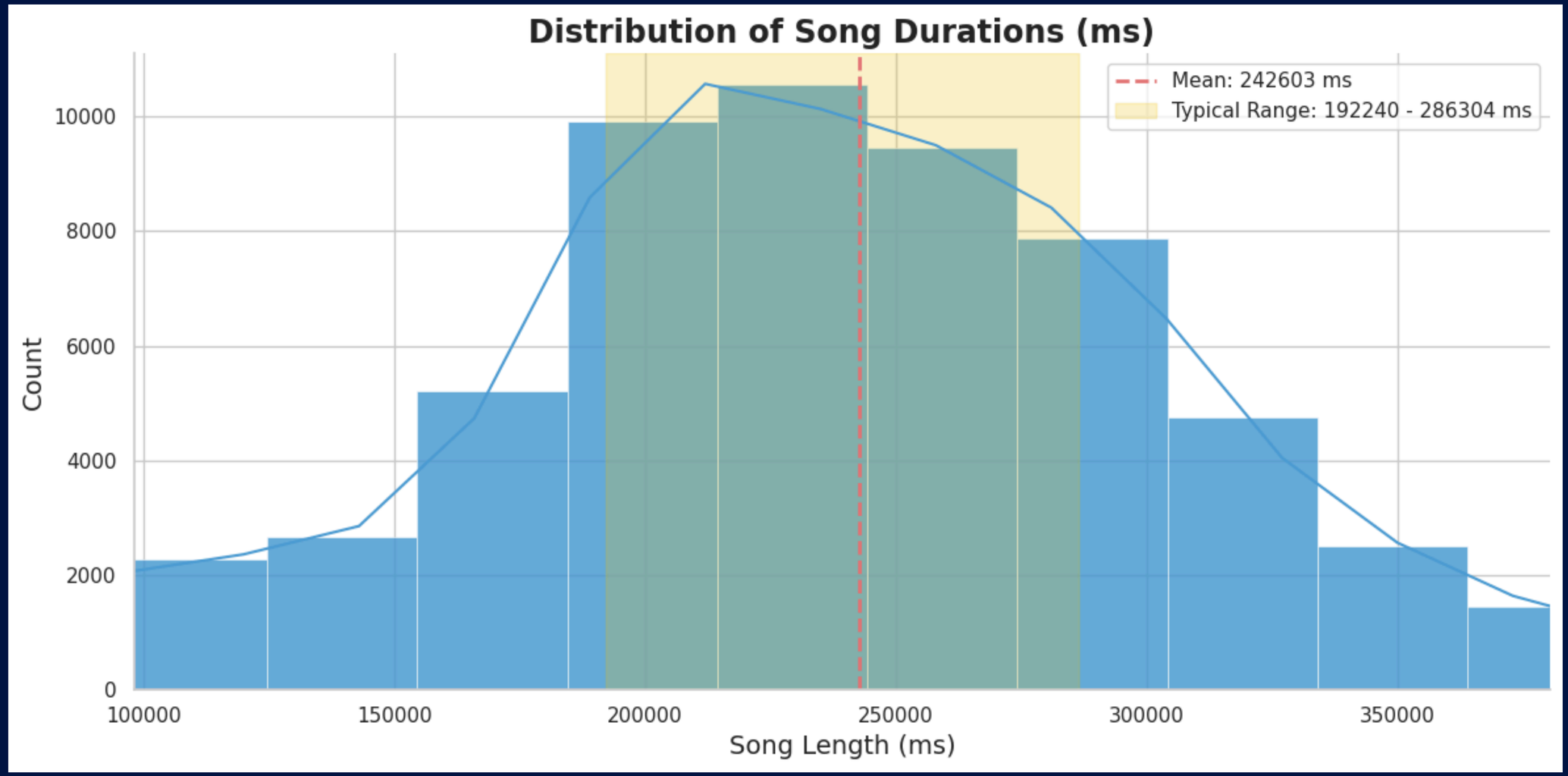
- Low (0–30): 79.15% of tracks
- Medium (31–70): 19.95% of tracks
- High (71–100): Only 0.89% of tracks

Insights:

- Most tracks have very low popularity (<30).
- The median (approx 7) and mean (approx 15) are both on the low side.
- The distribution is right-skewed, with a long tail toward high popularity.
- Only a small fraction of songs reach high popularity (above 70).
- The boxplot highlights many outliers at the higher end, representing hit songs.



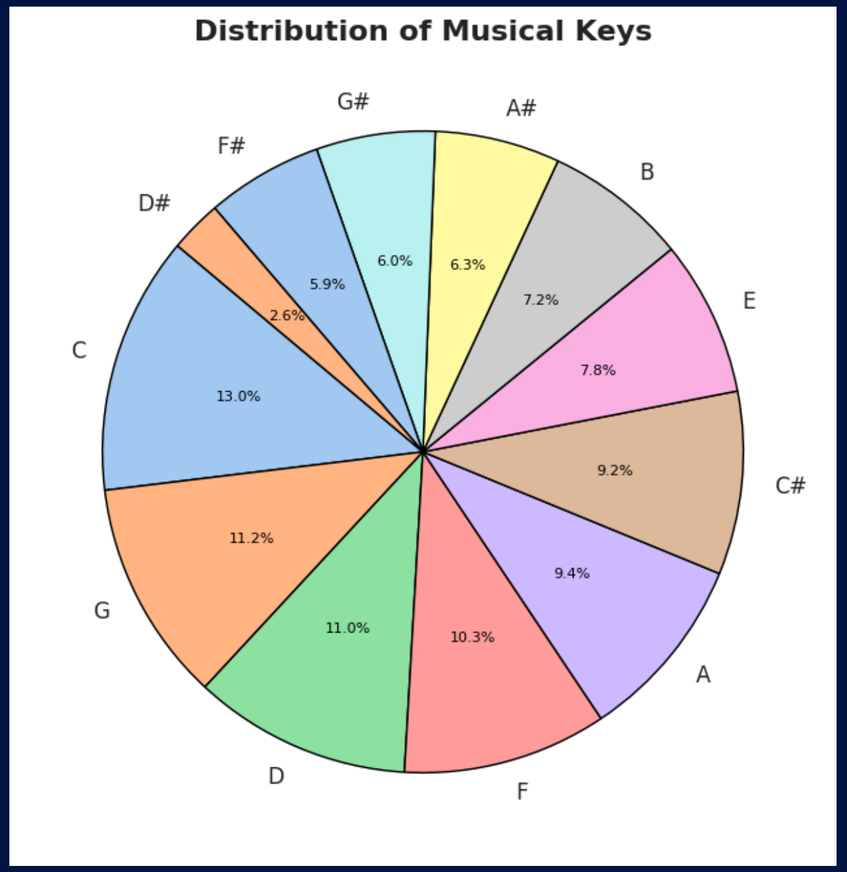
2. WHAT IS THE AVERAGE AND TYPICAL RANGE FOR DURATION_MS (SONG LENGTH)?

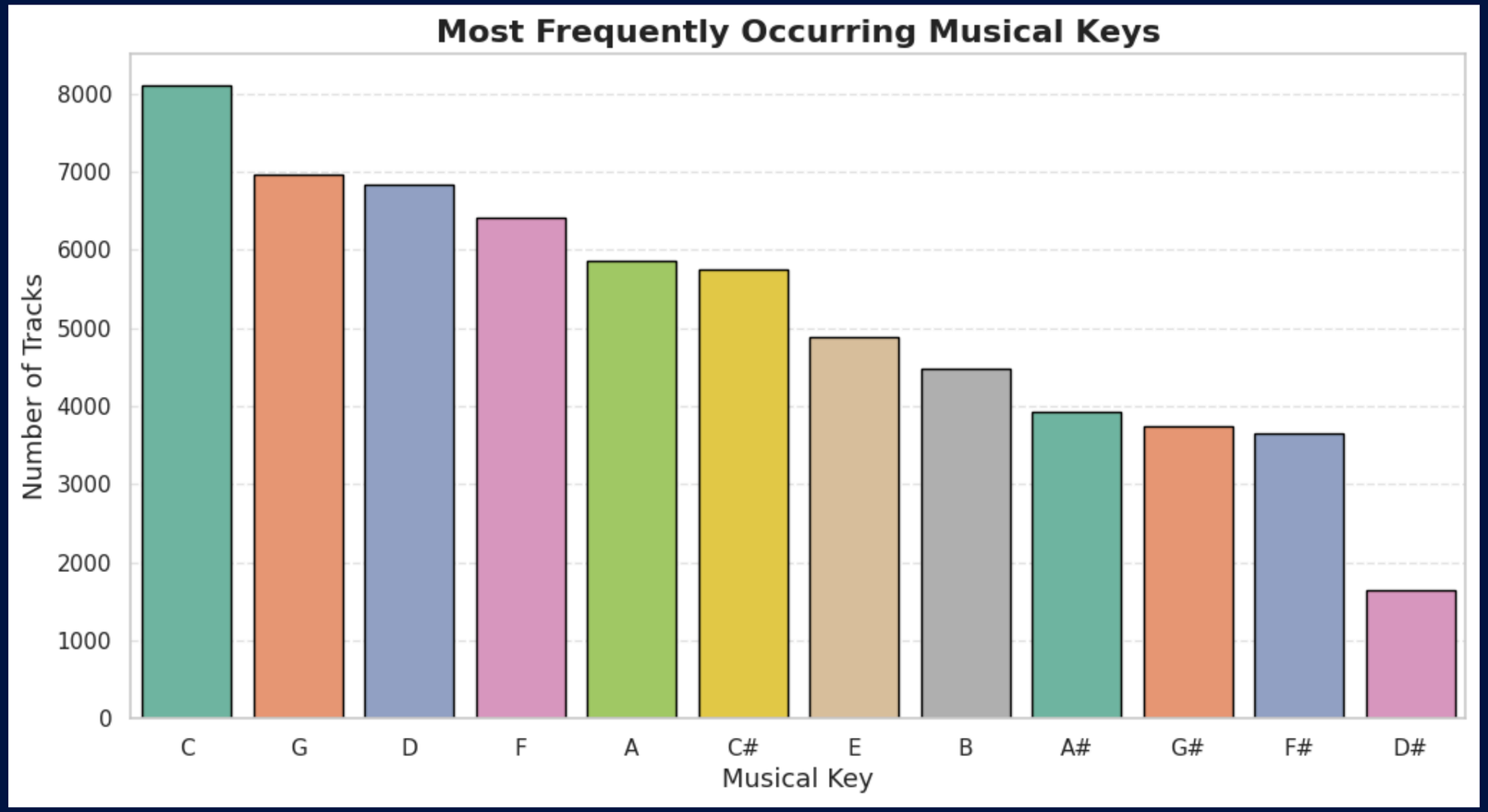


Insights:

- Average song duration: 242603.45 ms
- Typical range (25th to 75th percentile): 192240.00 ms to 286303.50 ms

3. WHAT ARE THE MOST FREQUENTLY OCCURRING KEYS IN THE DATASET, AND WHAT IS THEIR INDIVIDUAL DISTRIBUTION?





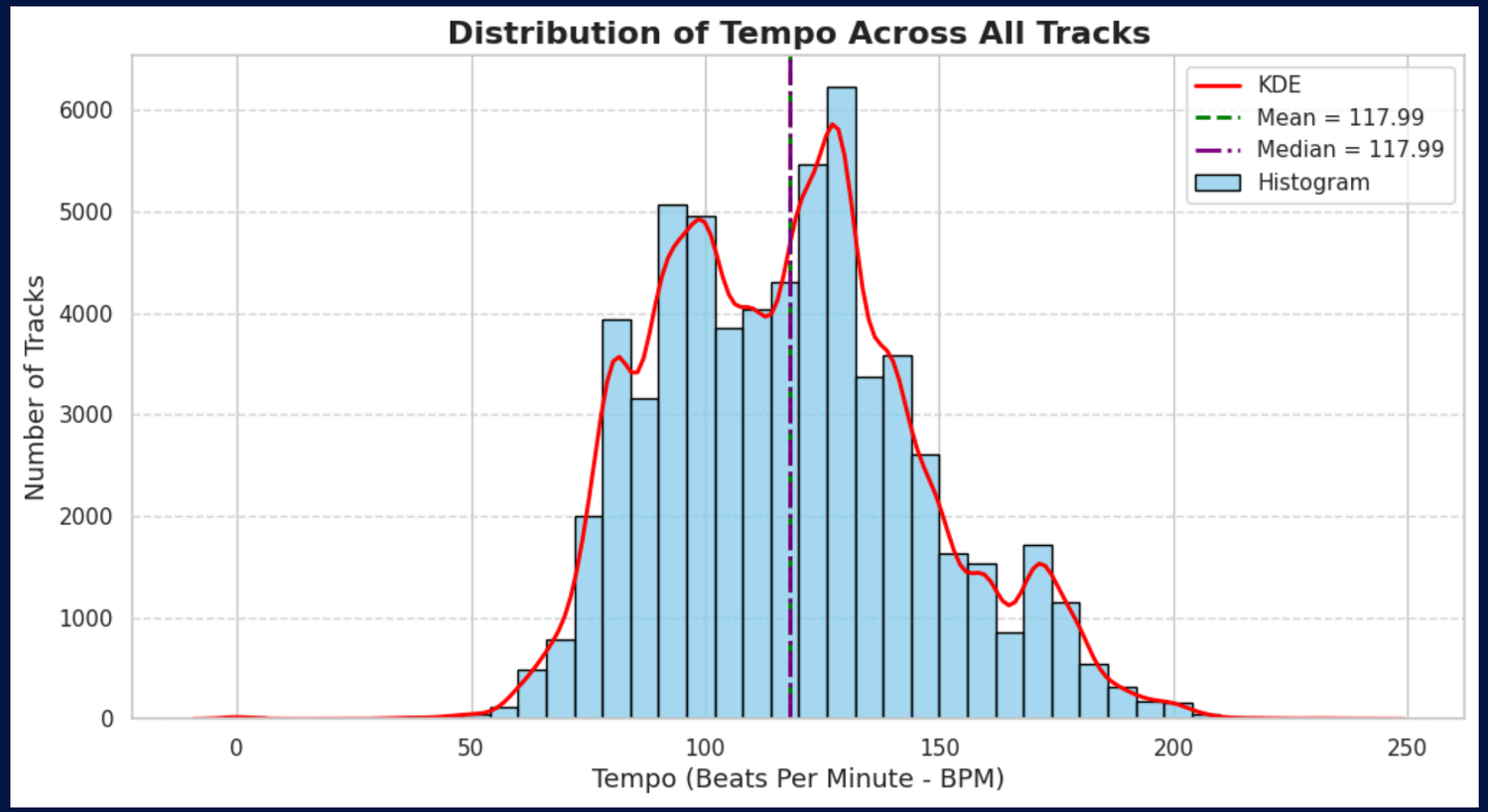
Key Frequencies:-

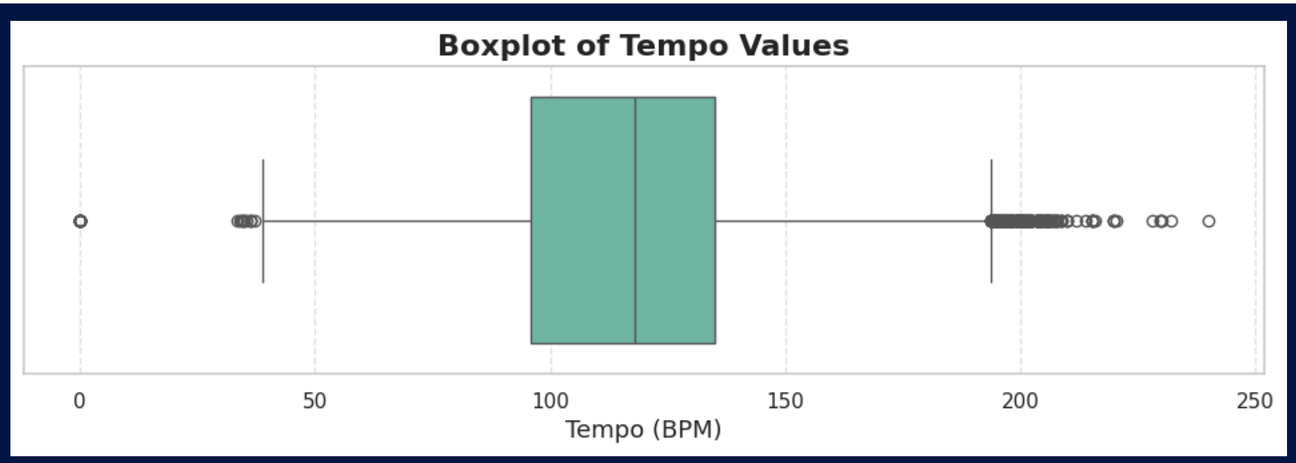
C: 8101	A: 5862	A#: 3920
G: 6956	C#: 5744	G#: 3735
D: 6832	E: 4879	F#: 3655
F: 6411	B: 4473	D#: 1636

Insights:

- Keys C, G, and D are the most frequently used, showing their dominance in popular music.
- Keys like C# and G# are less common, likely due to being harder to play or less conventional.
- The overall distribution highlights a preference for simple and versatile keys in mainstream tracks.

4. HOW ARE TEMPO VALUES DISTRIBUTED ACROSS ALL TRACKS? (ARE SONGS GENERALLY FAST, SLOW, OR IS THERE A WIDE SPREAD?)

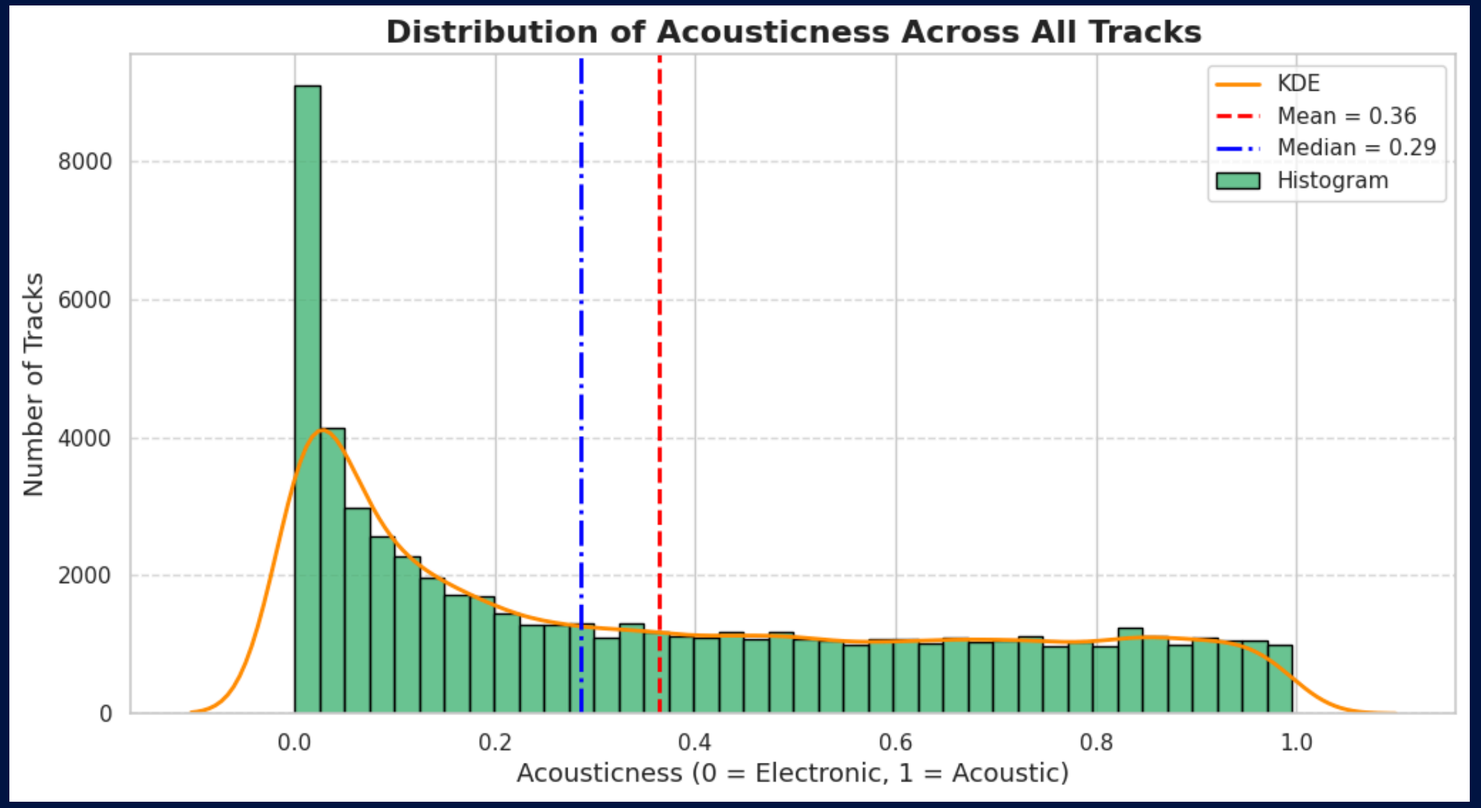


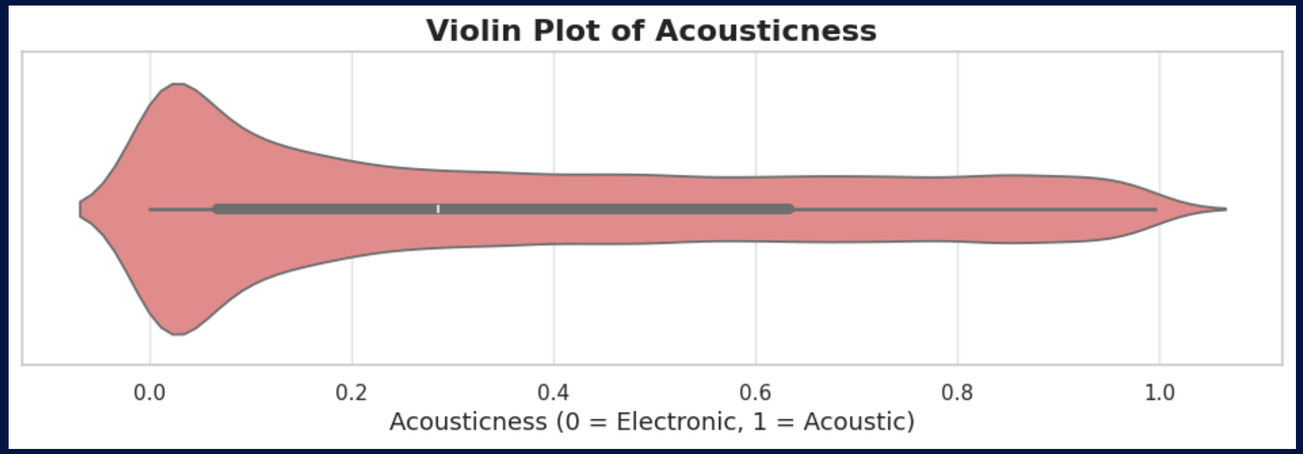


Insights:

- The distribution will likely show a wide spread (songs ranging from very slow (< 70 BPM) to very fast (>180 BPM)).
- Most tracks generally cluster around 100–130 BPM, a common tempo for pop, dance, and hip-hop songs.
- There are fewer songs at the extremes (very slow or very fast).
- Mean and median lines help compare whether tempo distribution is symmetric or slightly skewed.

5. WHAT IS THE DISTRIBUTION OF ACOUSTICNESS SCORES? (DOES THE DATASET LEAN TOWARDS ACOUSTIC OR ELECTRONIC SOUNDS?)

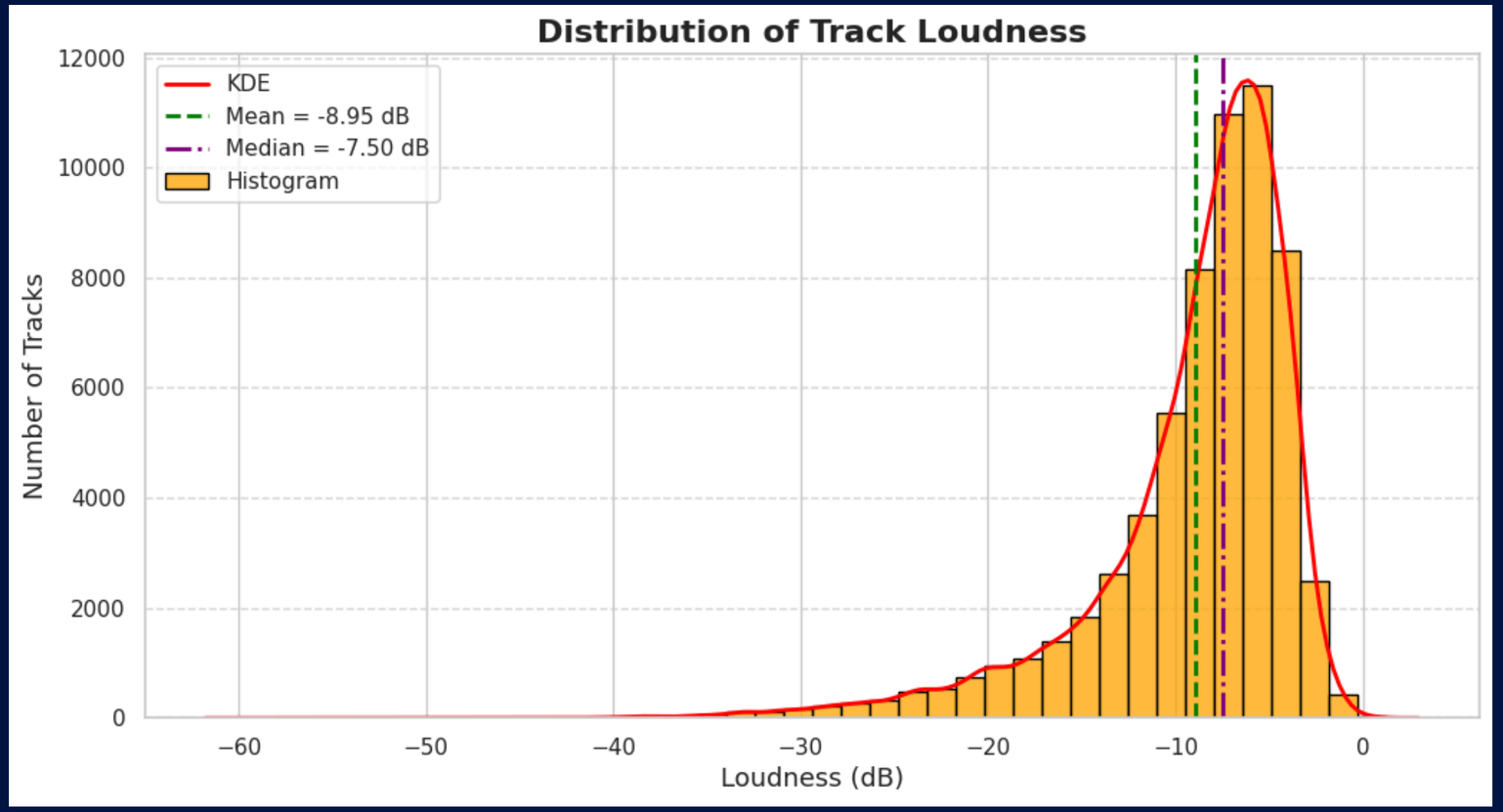


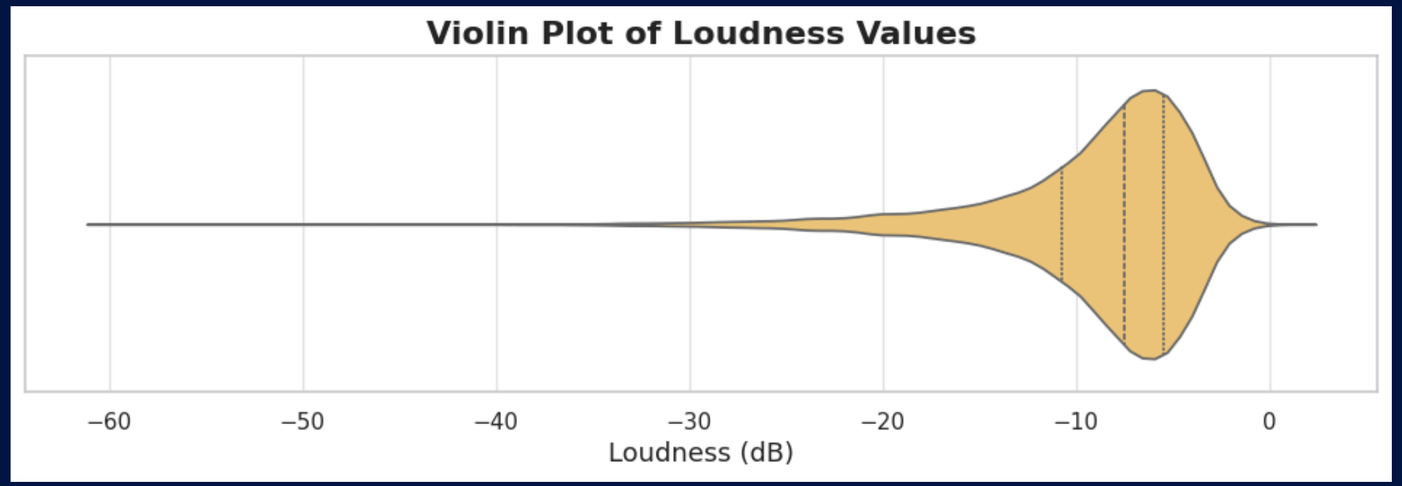


Insights:

- The distribution is usually skewed towards 0–0.3, meaning most tracks are more electronic/produced.
- A smaller but visible bump often exists near 0.8–1.0, representing highly acoustic tracks (folk, unplugged, classical, etc.).
- The dataset therefore shows a dominance of electronic/produced music, with a smaller share of fully acoustic sounds.
- The violinplot highlights where the bulk of tracks fall and whether there are long tails at the high-acoustic end.

6. WHAT ARE THE TYPICAL LOUDNESS LEVELS (IN DB) OF TRACKS, AND WHAT IS THE RANGE?

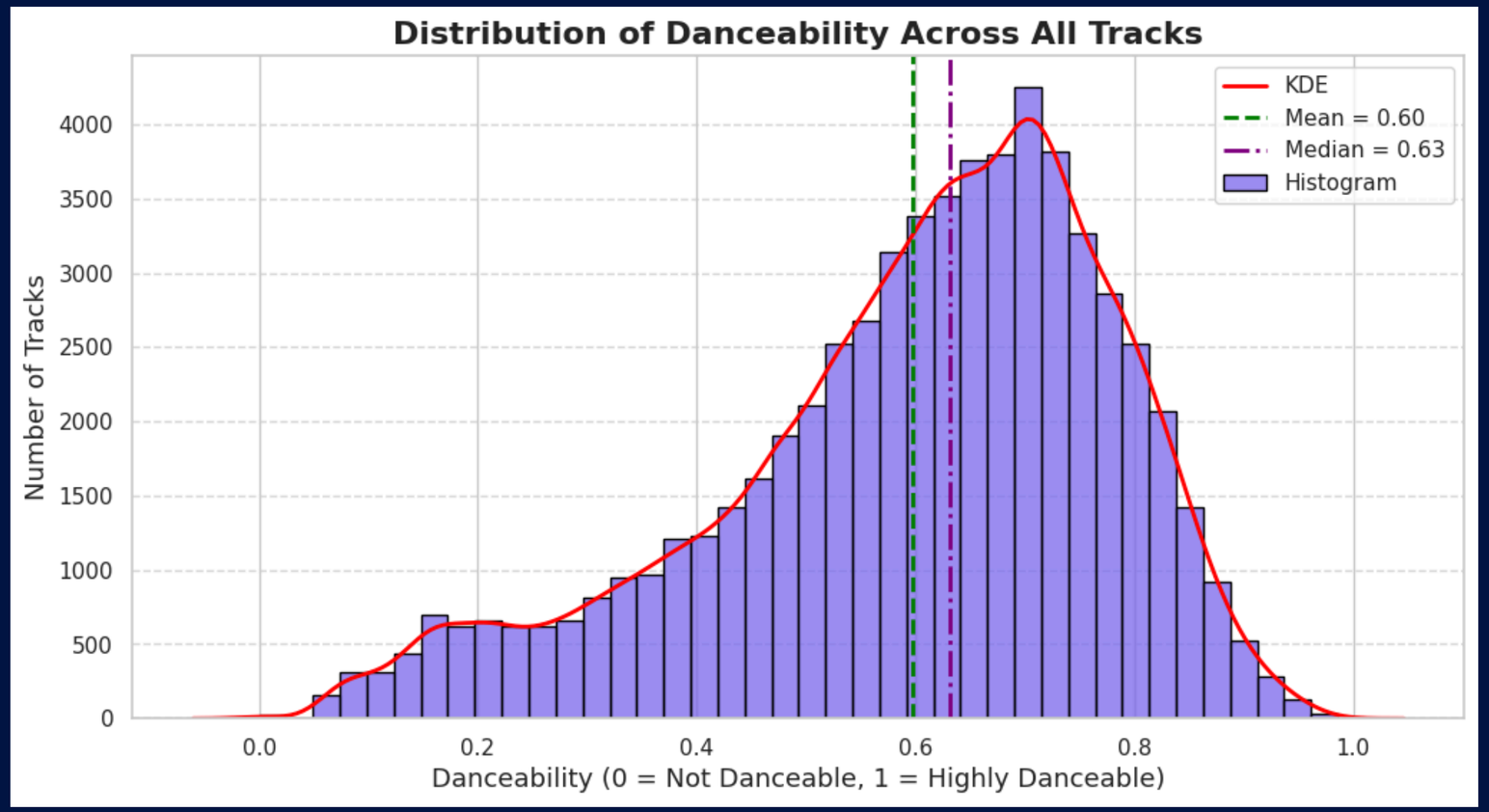


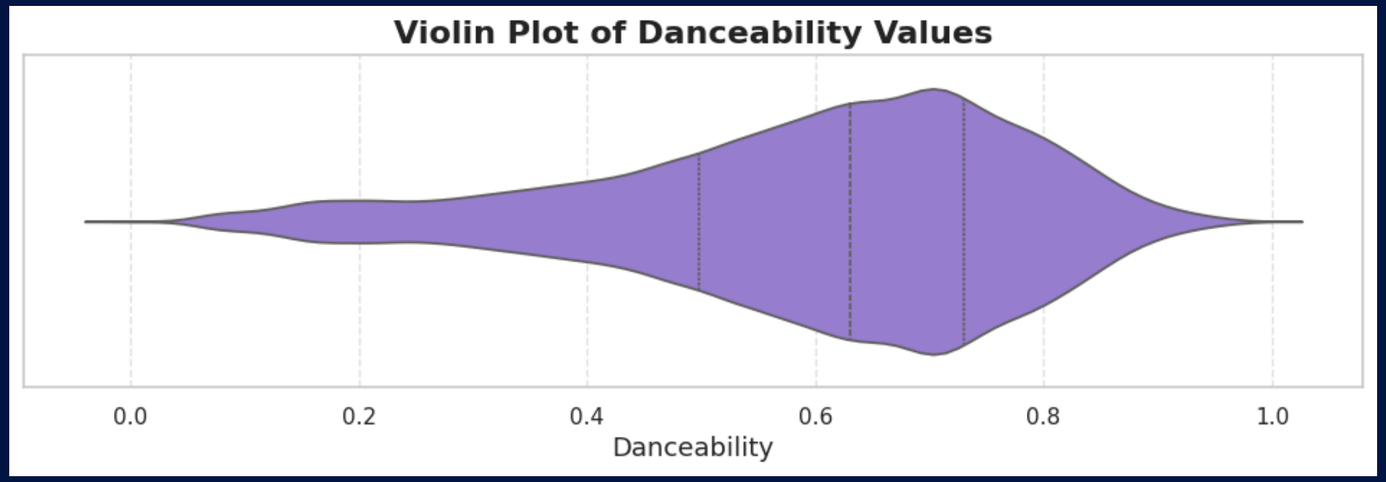


Insights:

- Typical Loudness: The typical loudness of a track is approximately -7.50 dB (median). The average is slightly lower at -8.95 dB due to a tail of quieter songs.
- Range: While the full range can span from below -60 dB up to 0 dB, the vast majority of tracks are concentrated in a much narrower range, typically between -15 dB and -5 dB.
- The distribution is strongly left-skewed, which means that while most songs are clustered at a high loudness level, there is a long tail of exceptionally quiet tracks that are statistical outliers.
- The sharp peak very close to 0 dB is a classic indicator of the 'loudness war' in music production, where the industry trend is to master songs at the maximum possible volume to stand out.

7. HOW IS DANCEABILITY DISTRIBUTED? (ARE MOST SONGS HIGHLY DANCEABLE, OR IS THERE A MIX?)

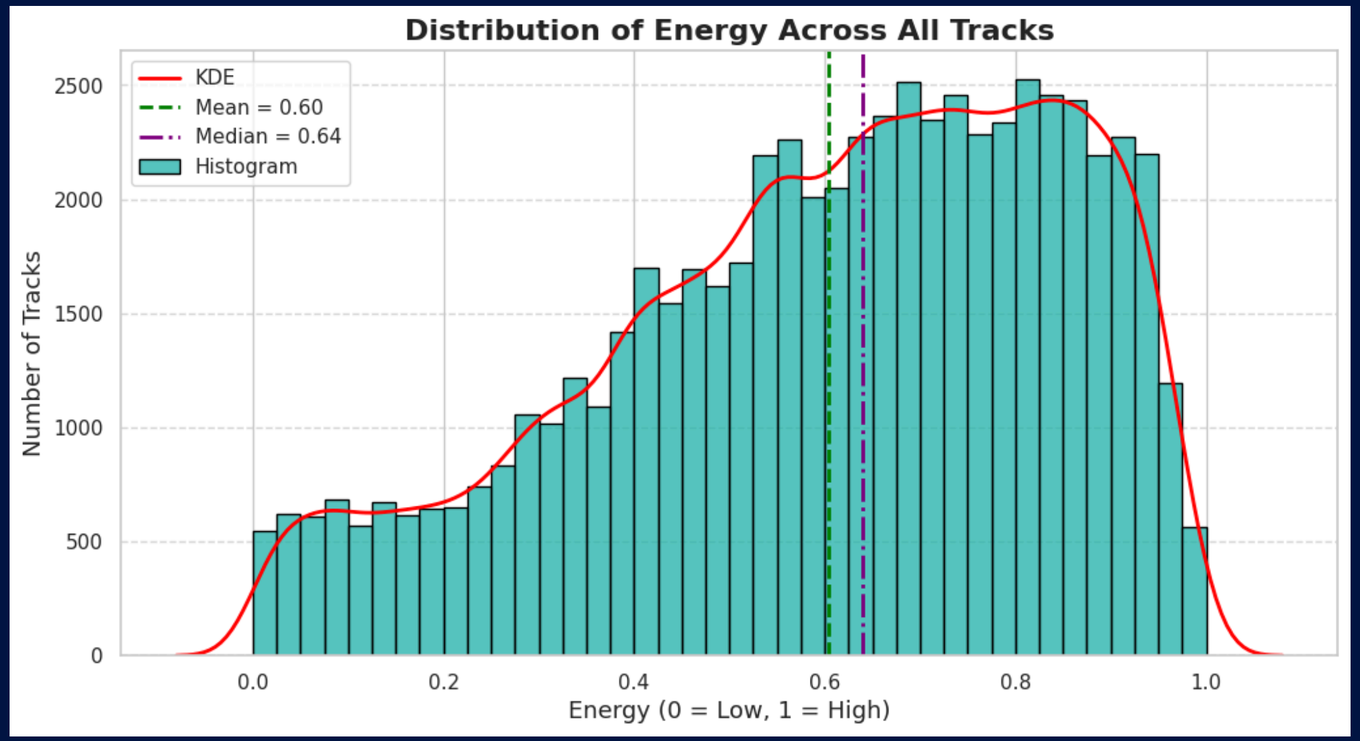


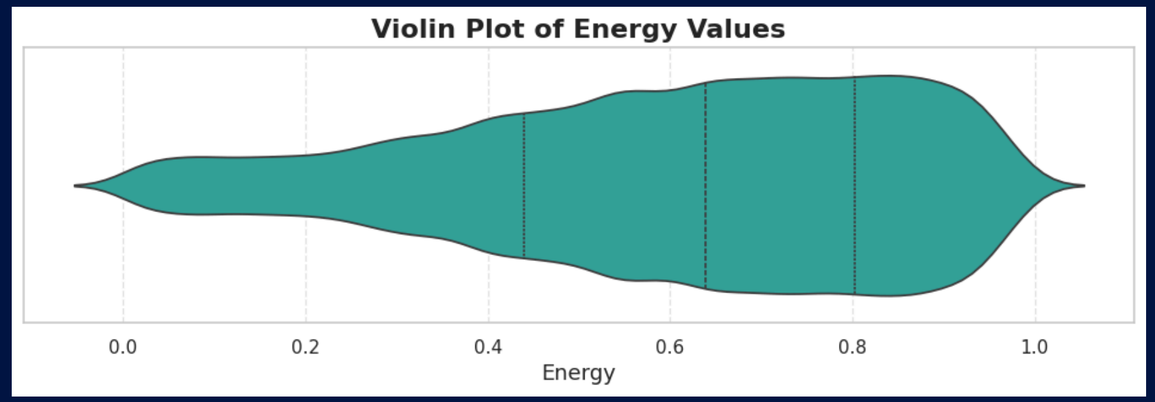


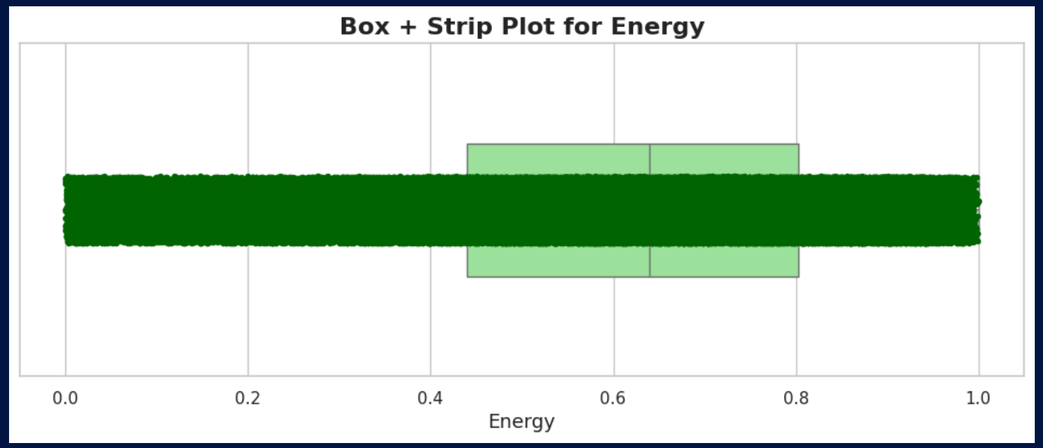
Insights:

- Danceability scores range from 0 to 1, with most songs clustered around 0.6–0.8, indicating a majority are moderately to highly danceable.
- Mean and median values are close, suggesting a fairly symmetric distribution.
- Violin plot shows fewer tracks with very low danceability (<0.3), so most tracks are indeed danceable.

8. WHAT IS THE DISTRIBUTION OF ENERGY LEVELS IN THE DATASET? (ARE SONGS GENERALLY HIGH OR LOW ENERGY?)







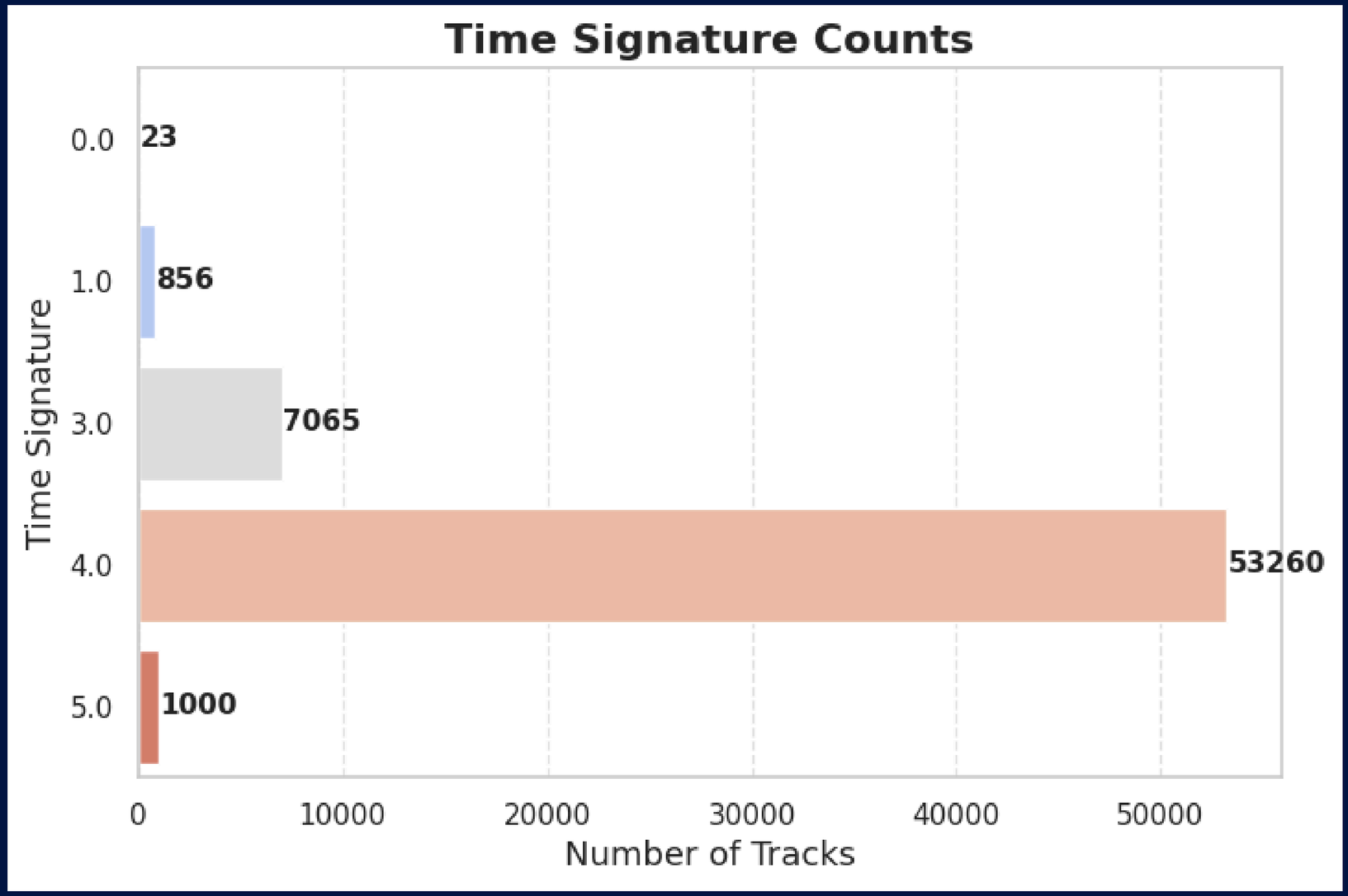
Insights:

- Typical Energy: The typical energy of a track is 0.64 (median) on a scale of 0 to 1. This indicates that a representative song from this catalog is moderately high in energy.
- Range: While the full range is from 0 (low) to 1 (high), the vast majority of tracks are concentrated in the upper half of the scale, primarily between 0.4 and 0.9.
- Sales Trends: The distribution is heavily left-skewed, revealing a strong catalog bias towards high-energy tracks. This suggests that the sales or streaming strategy is focused on upbeat music, which is often associated with popular genres like pop, EDM, and rock.
- Marketing & Promotions: The bimodal shape (a small peak at low energy and a large peak at high energy) indicates a clear segmentation in the catalog. This allows for targeted marketing.

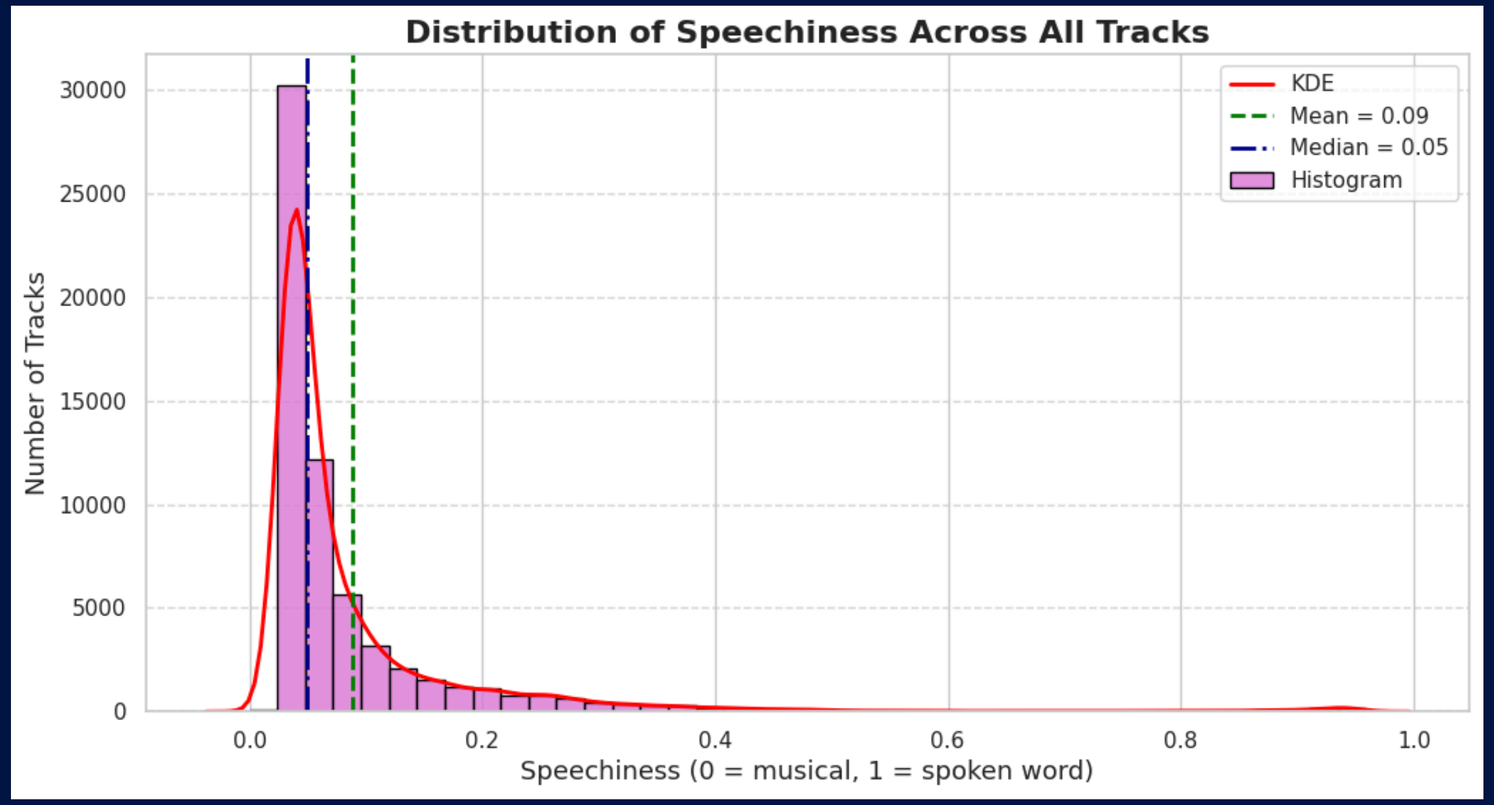
9. WHAT ARE THE MOST COMMON TIME_SIGNATURES FOUND IN THE MUSIC?

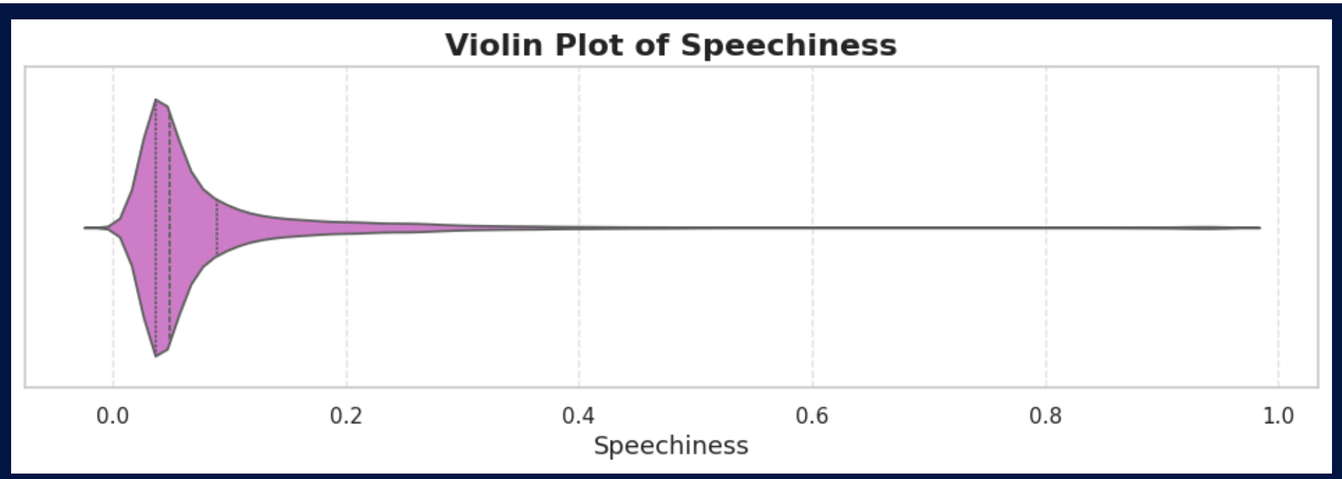
Insights:

- Total unique time signatures: 5
- Most common time signature: 4.0 (53260 tracks, 85.6% of dataset)
- Other time signatures proportions:
  - 3.0: 7065 tracks (11.4%)
  - 5.0: 1000 tracks (1.6%)
  - 1.0: 856 tracks (1.4%)
  - 0.0: 23 tracks (0.0%)



10. WHAT IS THE DISTRIBUTION OF SPEECHINESS? (ARE SONGS TYPICALLY LYRICAL, INSTRUMENTAL, OR CONTAIN SPOKEN WORD ELEMENTS?)





Insights:

- The dataset is overwhelmingly dominated by instrumental music, as the vast majority of tracks have a speechiness score near zero.
- While a few spoken-word tracks skew the average, the median (0.05) reveals the typical track is almost entirely musical, not lyrical.

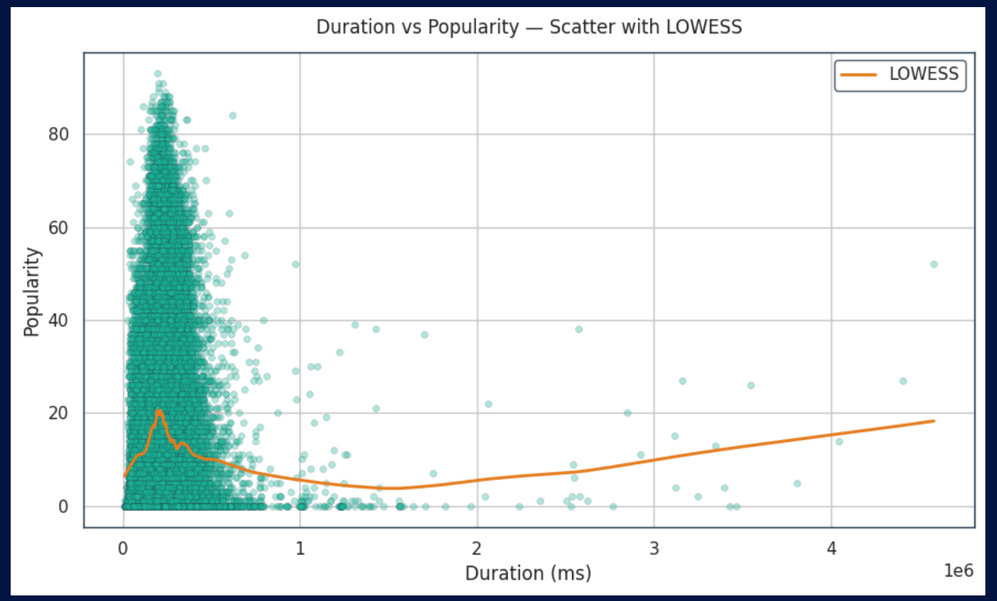
ANALYSIS AND KEY FINDINGS

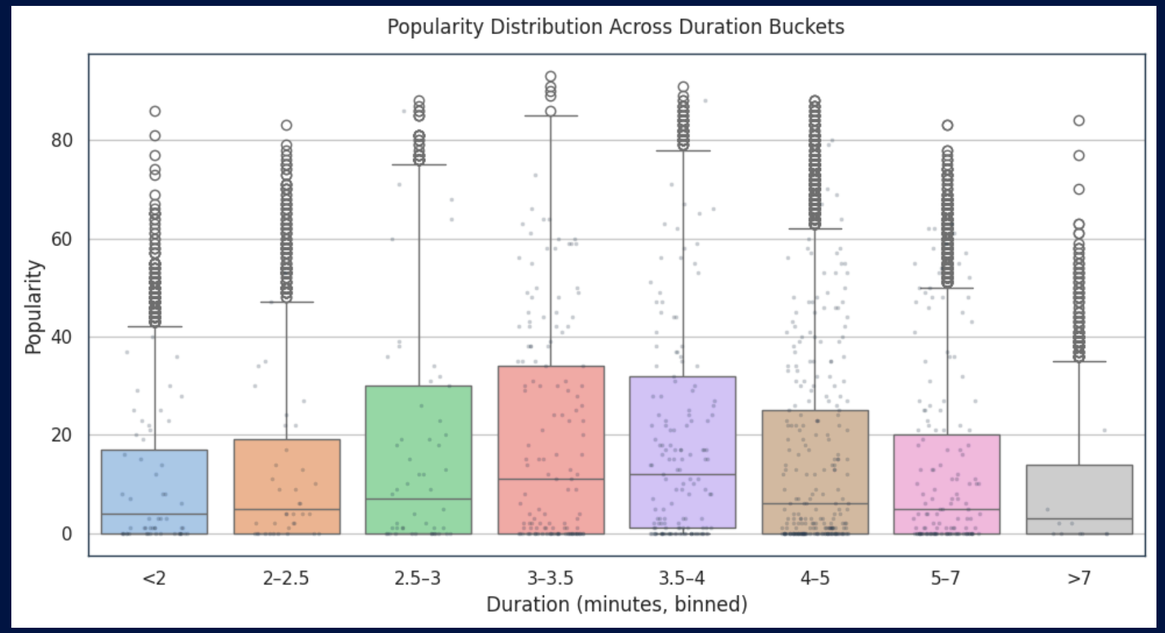
- Popularity: Most tracks fall in the moderate range, with fewer extremes.
- Duration: Songs are trending shorter (≈3 minutes today vs. longer in earlier years).
- Keys: C, G, and D dominate; D# and F# are least common.
- Tempo: Most songs lie in 90–120 BPM, an ideal groove range.
- Danceability: Popular songs skew high in danceability.
- Energy: Typically moderate-to-high, balancing intensity and engagement.
- Acousticness: Generally low → modern songs are more electronic.
- Loudness: Popular songs are louder, fitting modern mixing standards.
- Time Signatures: 4/4 is overwhelmingly dominant.
- Speechiness: Low for most tracks (mainly lyrical rather than spoken).

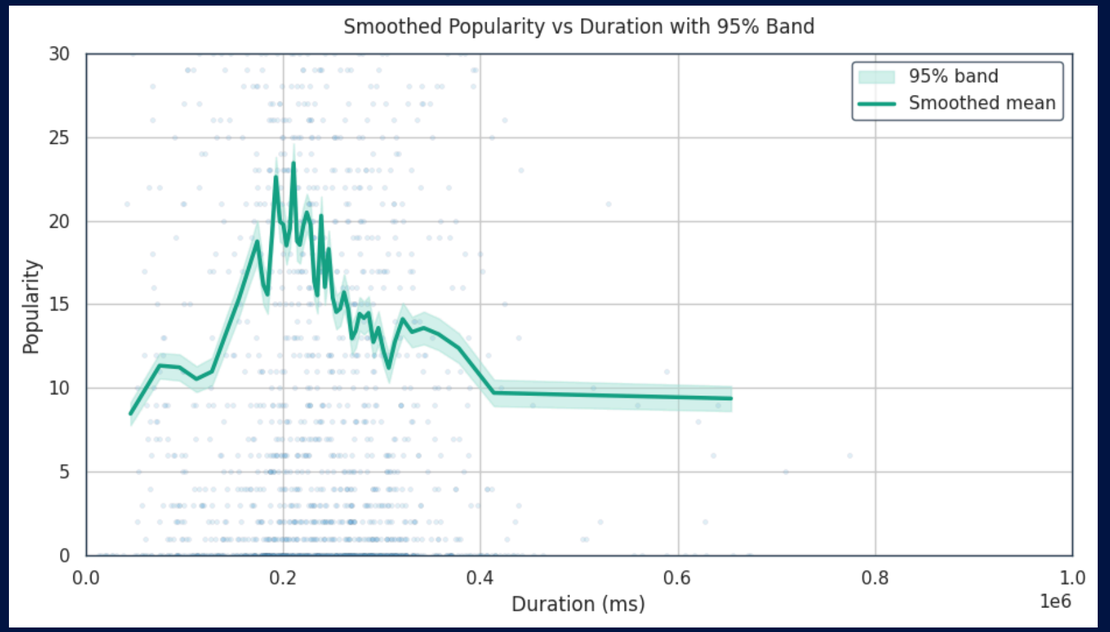
BIVARIATE ANALYSIS

Bivariate analysis explores the relationship between two variables to determine how they influence or correlate with each other. In this project, it helps uncover how musical features—such as danceability, energy, loudness, and duration—affect a song's popularity. By comparing pairs of attributes, this analysis reveals which factors contribute most strongly to making a track successful or appealing to listeners.

1. IS THERE A CORRELATION BETWEEN A SONG'S DURATION_MS AND ITS POPULARITY? (ARE SHORTER OR LONGER SONGS MORE POPULAR?)



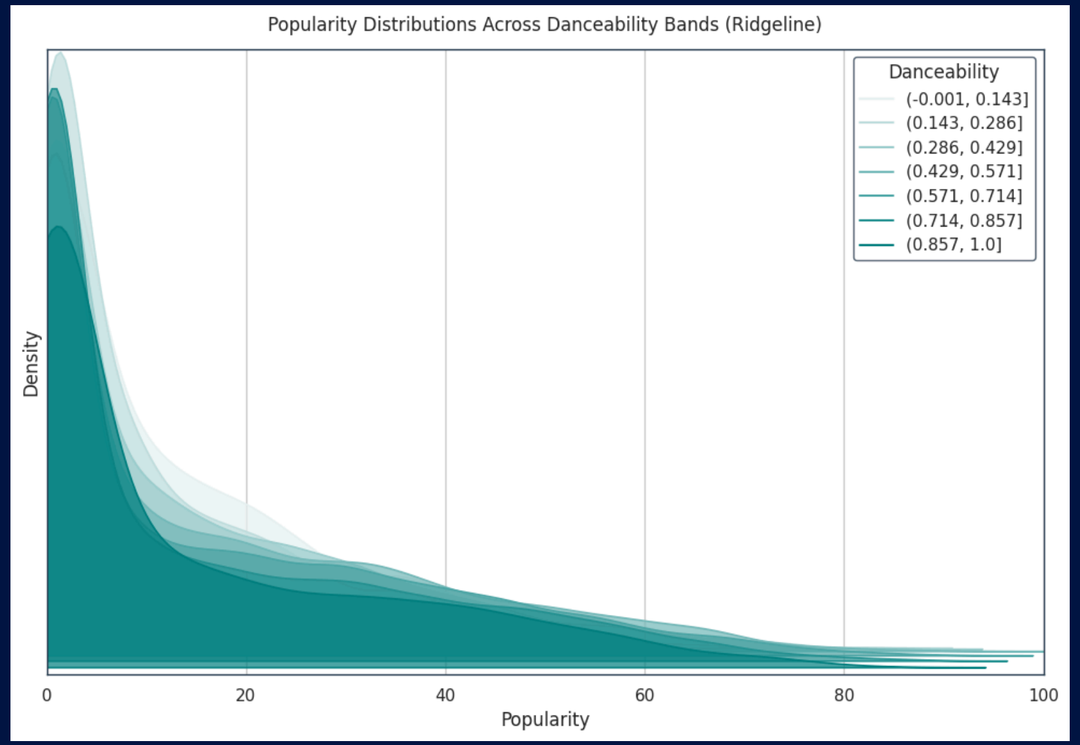




Insights:

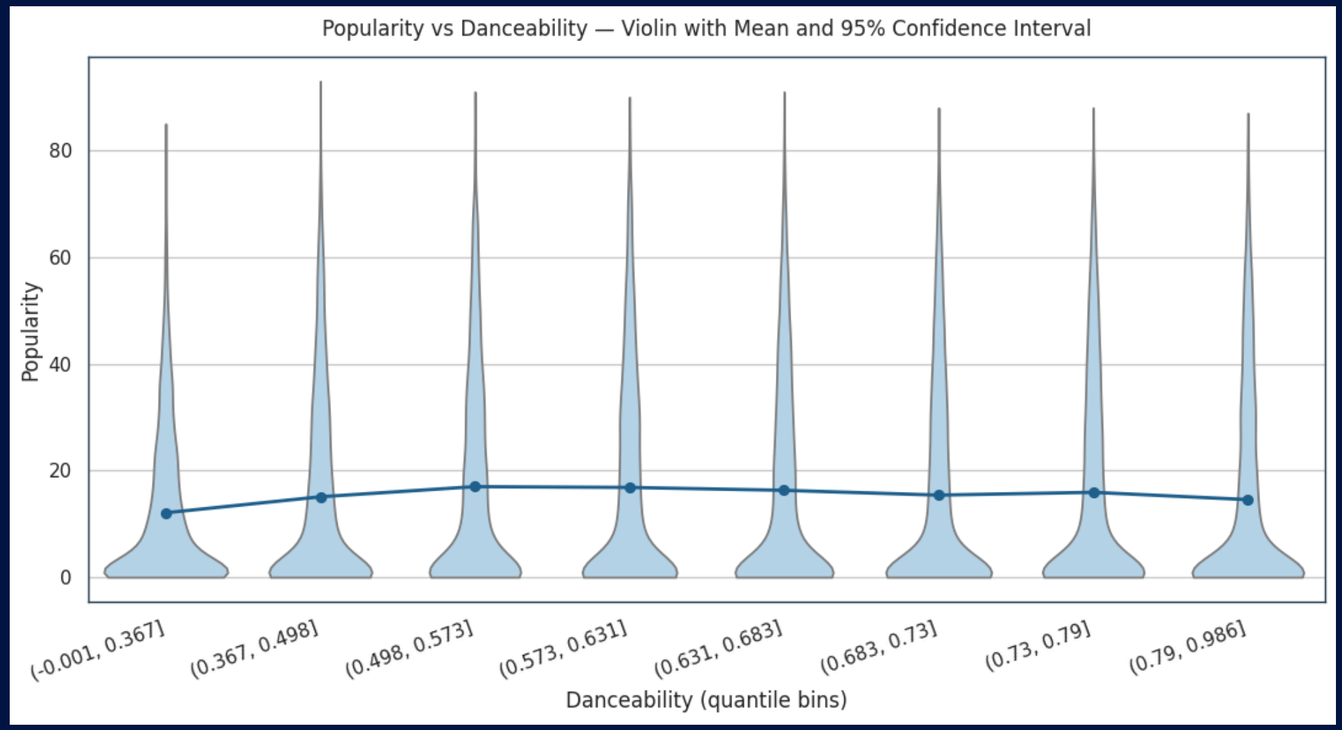
- The analysis shows that song popularity peaks between 200,000 and 350,000 milliseconds (about 3–6 minutes) — the optimal range for listener engagement. Popularity rises sharply around 3 minutes 20 seconds before gradually declining for longer tracks.
- Beyond the 4-minute mark, there's a clear drop in median popularity, indicating that shorter songs tend to perform better in mainstream music. The data is also heavily skewed, with a few hit songs significantly outperforming most others. Interestingly, a small number of very long tracks show a minor resurgence in popularity, likely from niche or extended versions.
- Overall, tracks around 3–4 minutes achieve the best balance between popularity and replay value.

2. HOW DOES DANCEABILITY RELATE TO POPULARITY? (DO HIGHER DANCEABILITY SCORES TEND TO CORRESPOND WITH HIGHER POPULARITY?)



Insights:

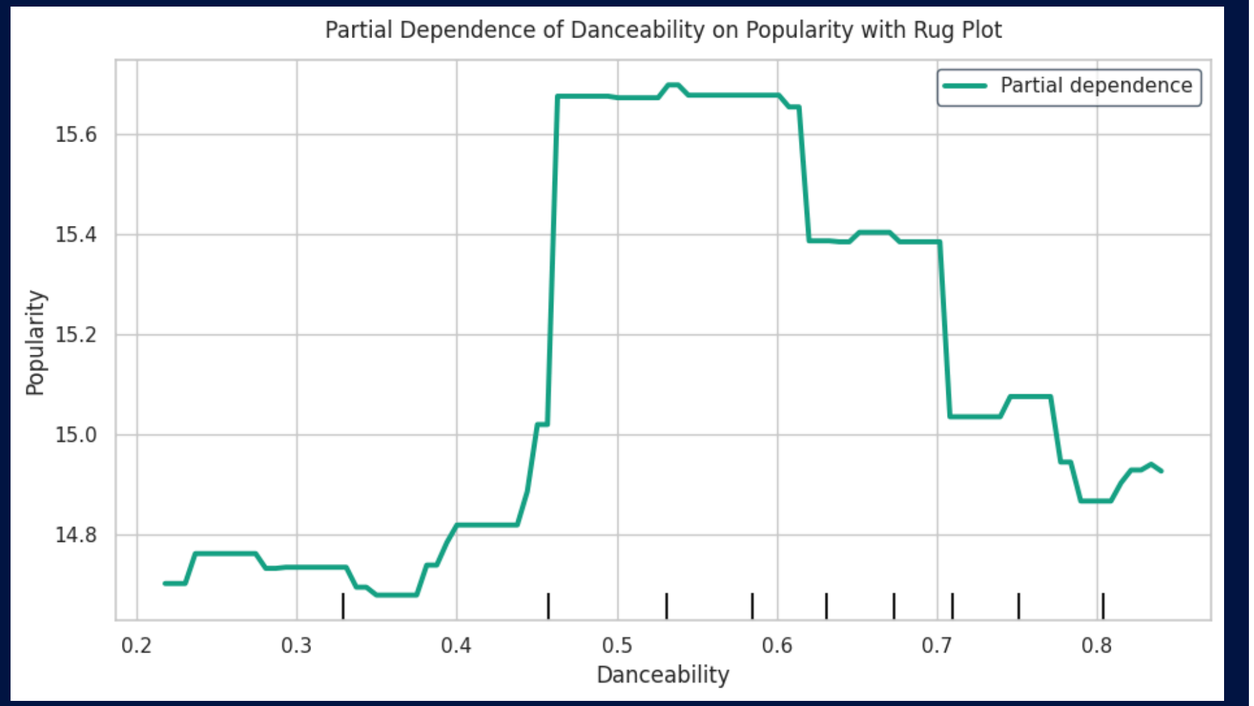
- All density curves peak at the same low popularity score. There is no evidence that the distribution for more danceable songs (darker colors) shifts towards higher popularity.
- The distributions for all danceability bands overlap almost completely, visually confirming that a song's popularity is largely independent of its danceability score.



Insights:

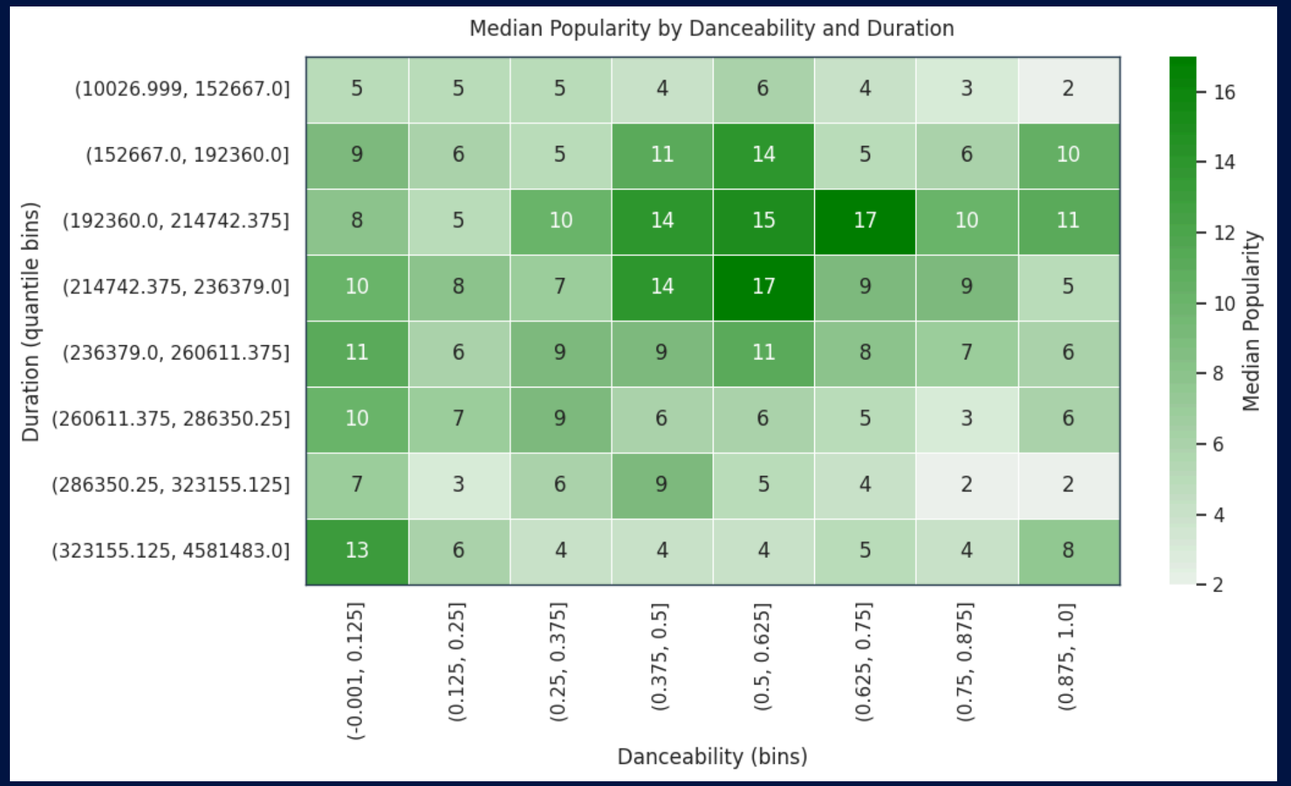
- The mean popularity (the connecting line) is almost perfectly flat across all danceability levels, showing that the average popularity of a song does not change with its danceability.
- The shape of the popularity distribution (the violin) is nearly identical for every bin, indicating that the probability of a song becoming a hit is the same regardless of whether its danceability is low or high.

2. HOW DOES DANCEABILITY RELATE TO POPULARITY? (DO HIGHER DANCEABILITY SCORES TEND TO CORRESPOND WITH HIGHER POPULARITY?)



Insights:

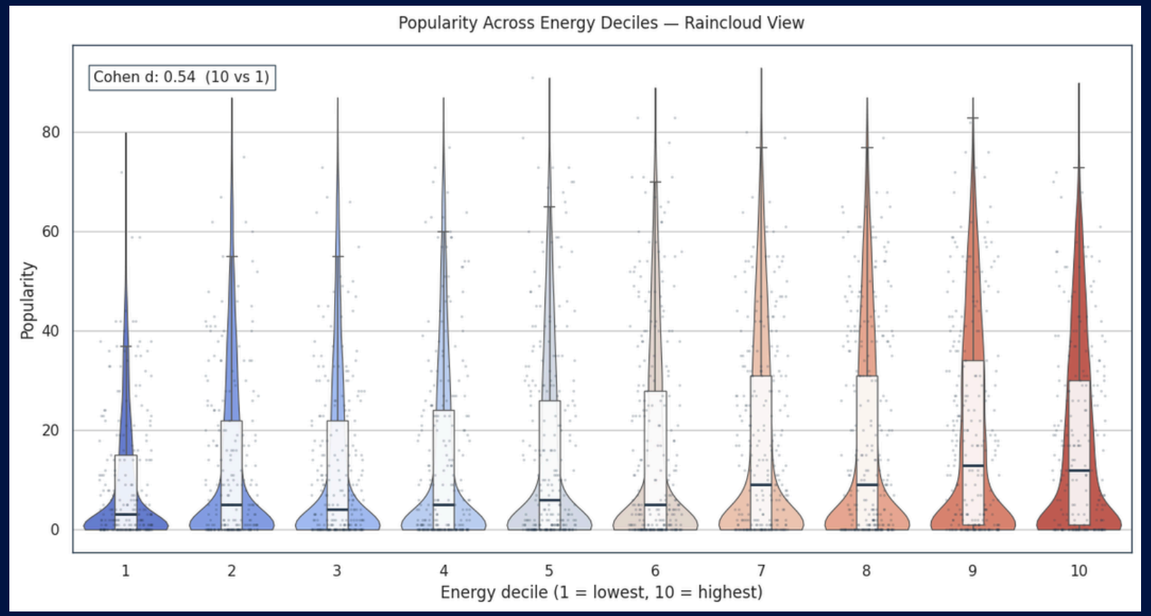
- The model's prediction for popularity hits a peak plateau for danceability scores between roughly 0.45 and 0.6, and does not continue to rise for higher scores.
- For danceability scores above 0.6, the model actually predicts a lower popularity, suggesting that being 'too danceable' can have a negative impact on a song's expected popularity.



Insights:

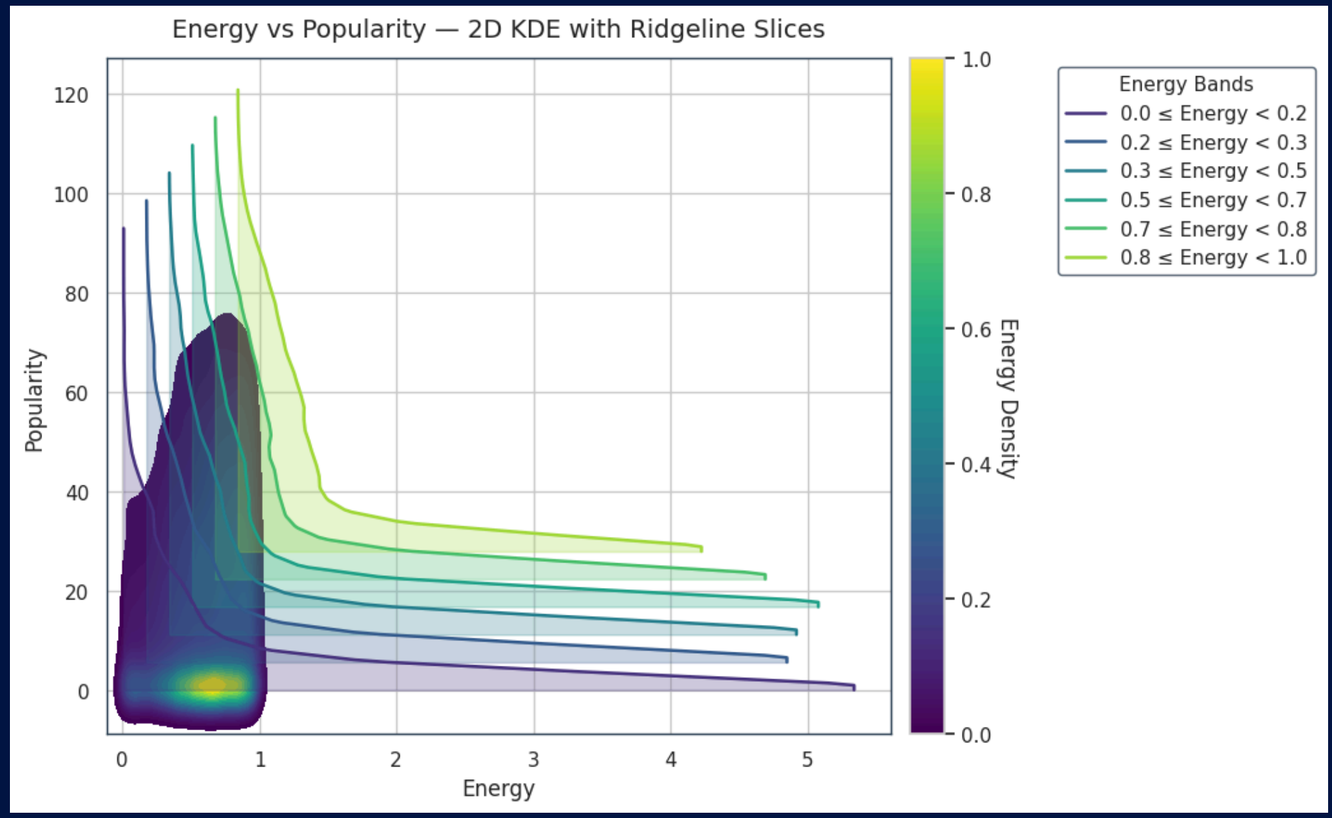
- The highest median popularity (darkest green cells) is found in the middle of the danceability axis (0.5-0.75), not at the extreme high end, which contradicts the idea that more danceable is more popular.
- Looking across any single row (i.e., for a fixed song duration), the color does not consistently get darker to the right, proving that increasing danceability alone does not reliably increase median popularity.

3. WHAT IS THE RELATIONSHIP BETWEEN ENERGY AND POPULARITY? (ARE HIGH-ENERGY TRACKS GENERALLY MORE POPULAR THAN LOW-ENERGY ONES?)



Insights:

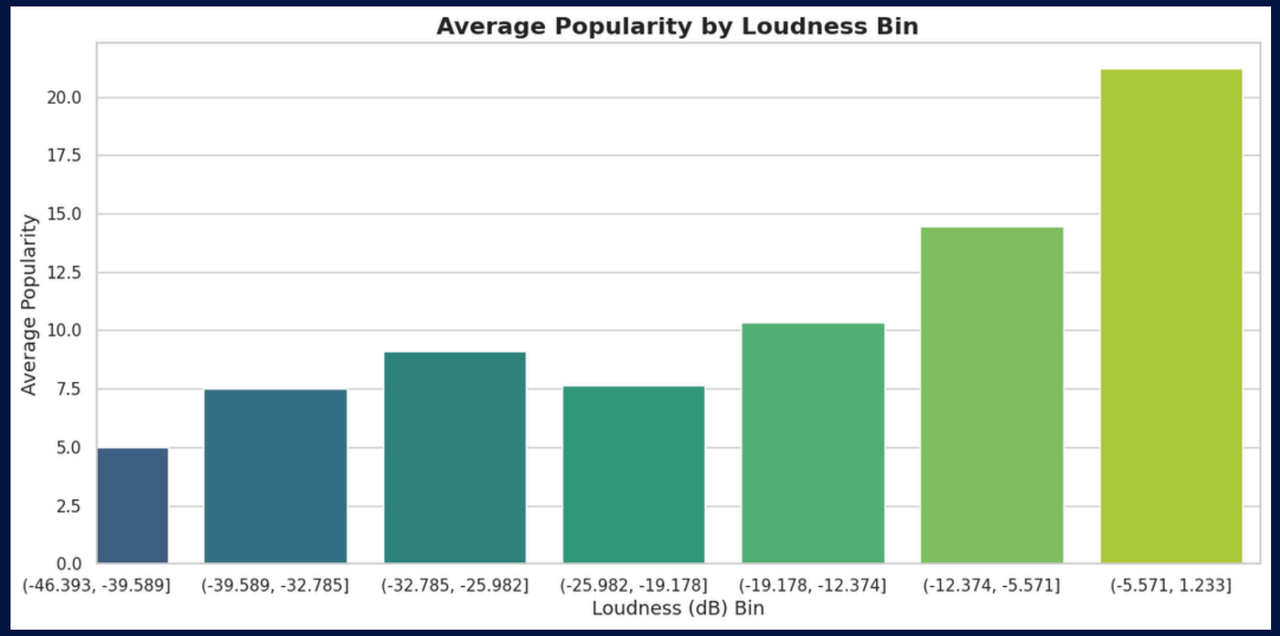
- There is a distinct and consistent rise in average popularity as you move from the lowest energy decile (1) to the highest (10). The entire distribution, including the median (white line), shifts upward.
- The "Cohen d: 0.54" indicates a medium-sized, statistically significant effect. This confirms that the observed increase in popularity for high-energy tracks over low-energy ones is not due to random chance.



Insights:

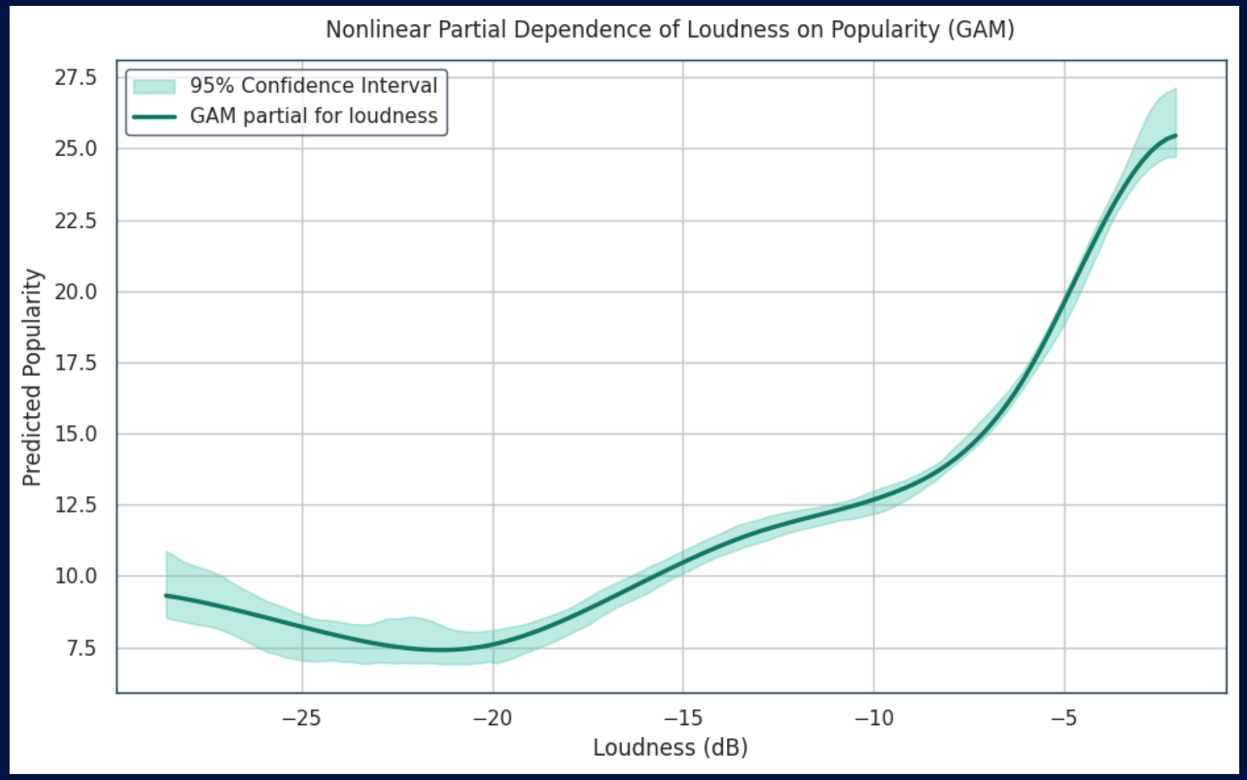
- The densest area of the plot is in the bottom-left corner, showing that the vast majority of all tracks have both low energy and low popularity.
- The songs that do achieve high popularity (e.g., > 60) are almost exclusively found in the higher energy bands (the green and yellow distributions). Low-energy tracks (purple) rarely reach high popularity levels.

4. DOES LOUDNESS HAVE A NOTICEABLE IMPACT ON POPULARITY? (ARE LOUDER MIXES PREFERRED BY LISTENERS?)



Insights:

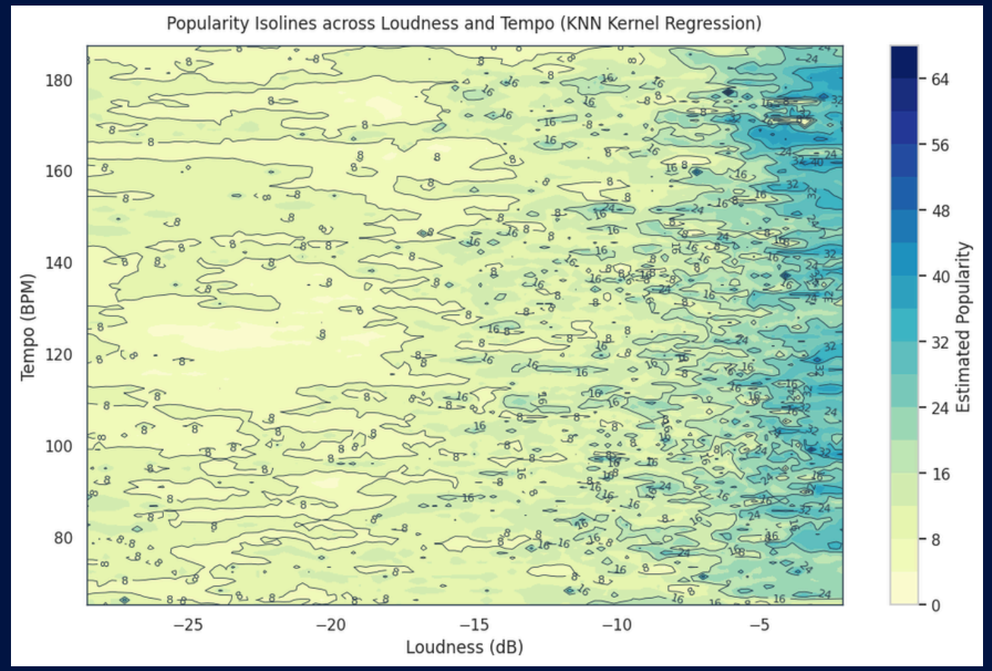
- The bar chart shows a clear, step-by-step increase in average popularity as the loudness bin gets higher. There isn't a single instance where a louder category of songs is less popular on average than a quieter one.
- The biggest jump in popularity occurs in the final bin (from -5.571 dB upwards). This suggests that while being loud is good, being very loud provides a disproportionately large boost to a track's average popularity.



Insights:

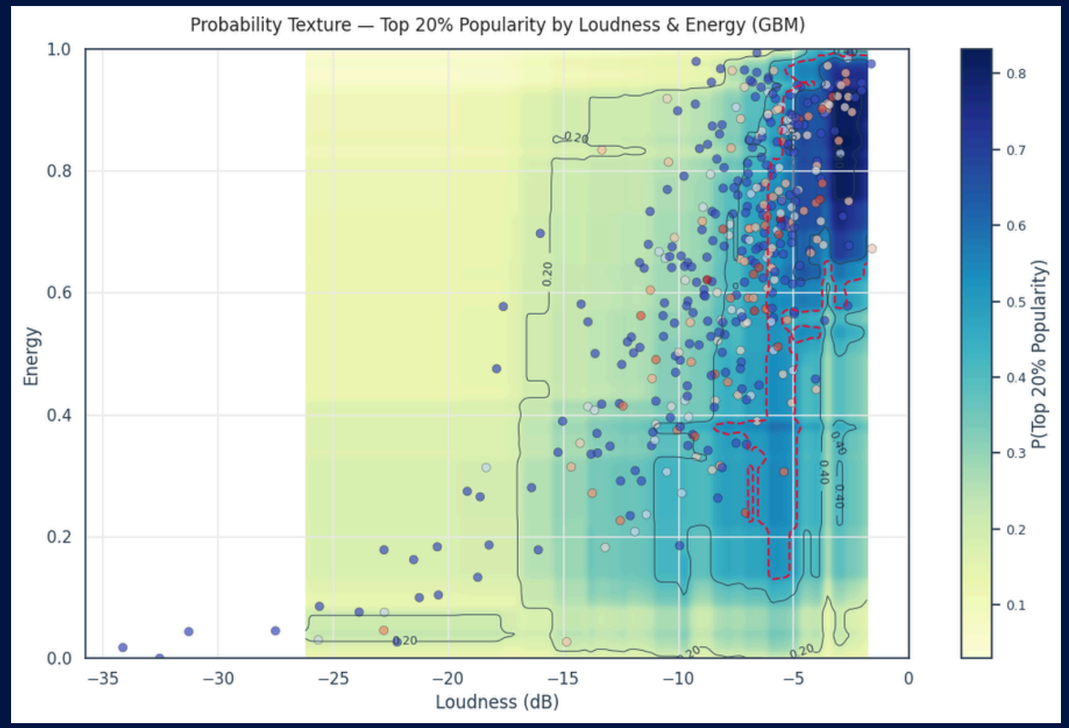
- The impact of loudness isn't linear. The curve is relatively flat for quiet songs (from -25 to -15 dB), meaning making a quiet song slightly less quiet does little. However, it becomes exponentially steeper after -10 dB, showing that increases in loudness for already-loud tracks have a massive impact on predicted popularity.
- The graph reveals a "penalty zone" below -20 dB where predicted popularity hits its lowest point. This indicates that mixes perceived as too quiet are actively disliked, more so than tracks in a more moderate loudness range.

4. DOES LOUDNESS HAVE A NOTICEABLE IMPACT ON POPULARITY? (ARE LOUDER MIXES PREFERRED BY LISTENERS?)



Insights:

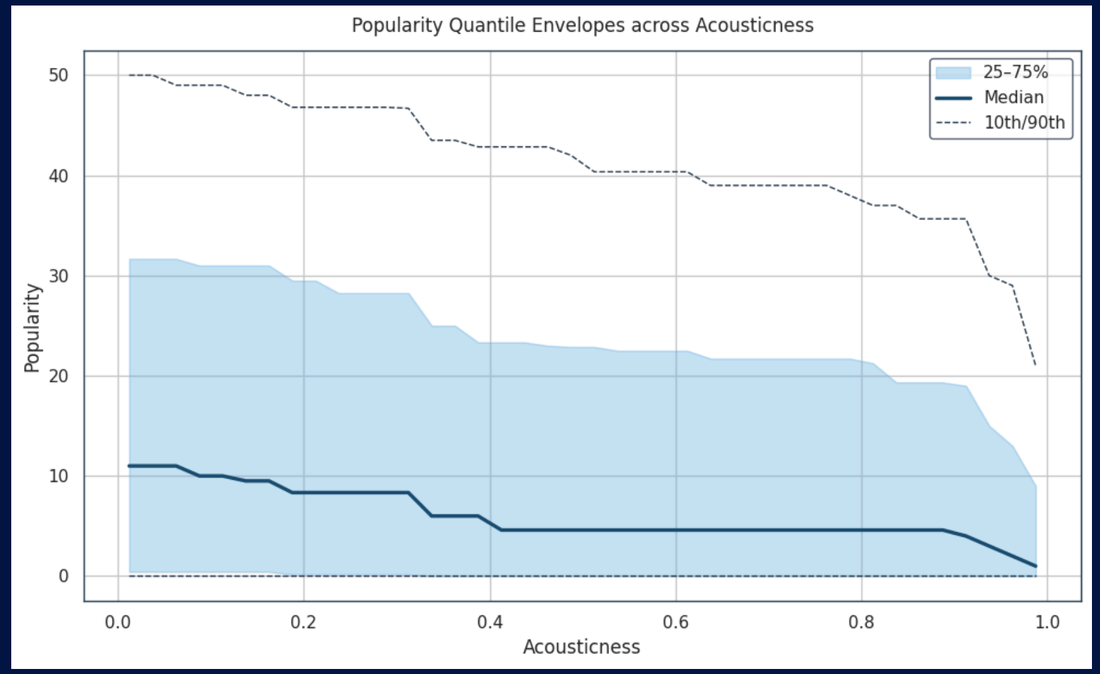
- Moving from left to right on the graph (increasing loudness) results in a much more dramatic increase in popularity (the color changing from yellow to dark blue) than moving up or down (changing tempo). This shows loudness is a more powerful driver of popularity than a song's speed.
- There is a clear "wall" around -8 dB to -10 dB. Tracks quieter than this (to the left) rarely achieve high popularity, while the vast majority of popular tracks (dark blue areas) are louder than this threshold.



Insights:

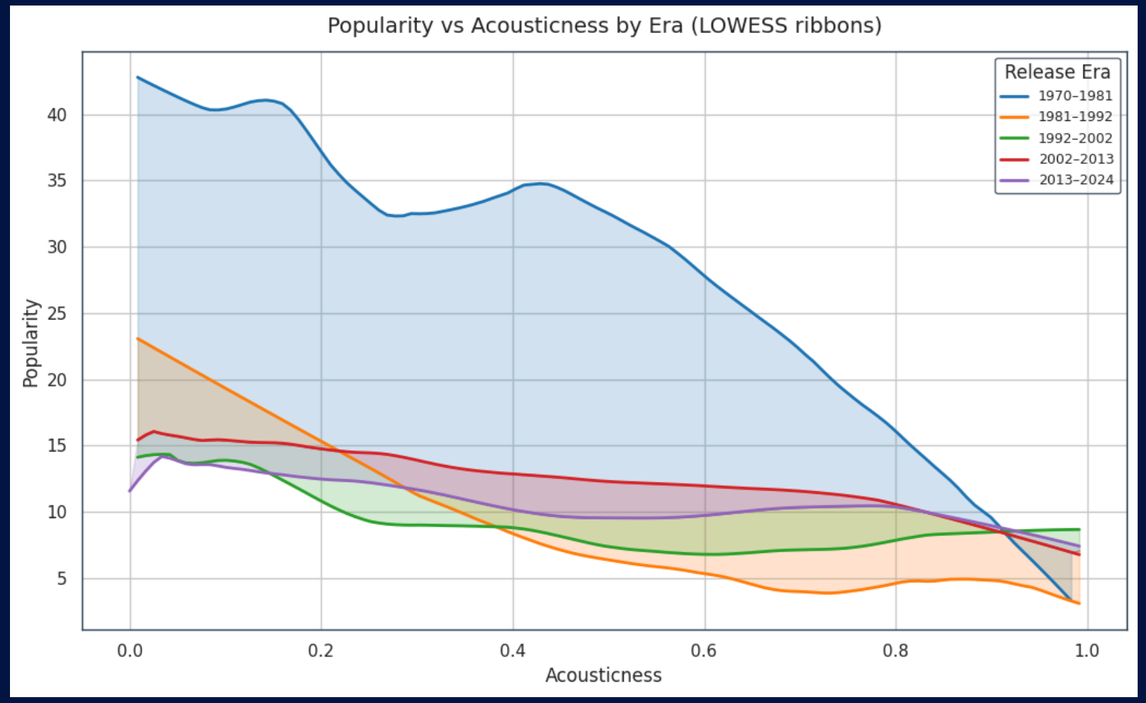
- High loudness is a necessary condition for a song to have a high probability of being a top-tier hit. Even tracks with maximum energy have virtually no chance (yellow/green color) of entering the top 20% of popularity if their loudness is below -15 dB.
- The highest probability of a song becoming a major hit (the dark blue region where P > 0.8) exists only in the top-right corner. This requires a combination of very high loudness (>-5 dB) and high energy (>0.8), making this the optimal zone for producing popular music.

5. IS THERE A RELATIONSHIP BETWEEN ACOUSTICNESS AND POPULARITY? (ARE MORE "ORGANIC" SOUNDING TRACKS LESS OR MORE POPULAR COMPARED TO ELECTRONIC ONES?)



Insights:

- The median popularity (the solid dark blue line) is highest for tracks with near-zero acousticness. It declines significantly as acousticness increases, showing that the average song is much more popular when it's not acoustic.
- The dashed line representing the 90th percentile (the most popular hits) starts very high (popularity of 50) for non-acoustic tracks and steadily decreases. This means the potential for a track to become a massive global hit is significantly greater if it is electronic/produced.



Insights:

- Every single line, representing different eras from the 1970s to the 2020s, shows the same downward trend. This preference for non-acoustic music is not a recent phenomenon but a stable pattern that has persisted for over 50 years.
- While less popular overall, the popularity of highly acoustic tracks (acousticness > 0.8) has remained remarkably consistent across all decades. This suggests a dedicated, stable niche audience for "organic" music that exists outside of mainstream trends.

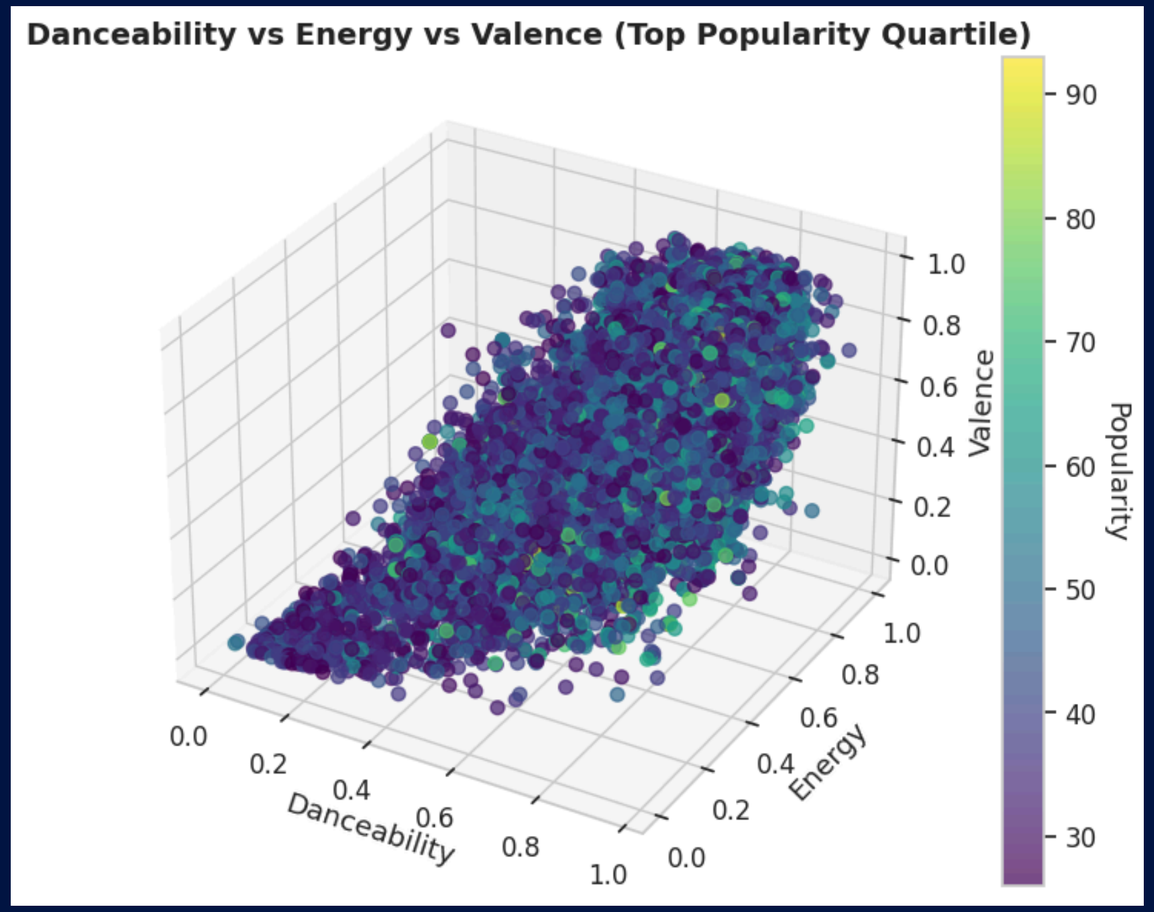
ANALYSIS AND KEY FINDINGS

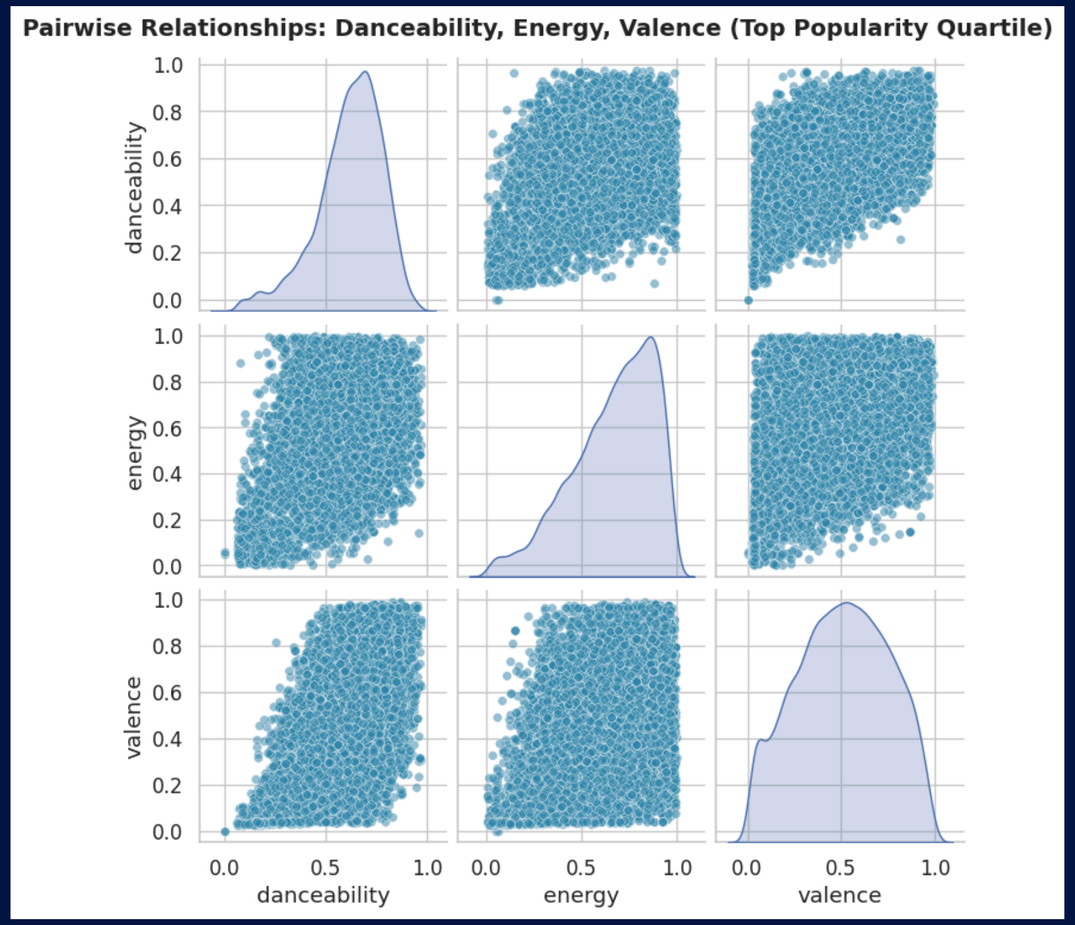
- Danceability ↔ Popularity: Stronger danceability = higher popularity.
- Energy ↔ Popularity: Balanced energy performs better than extremes.
- Loudness ↔ Popularity: Louder songs are slightly more popular.
- Duration ↔ Popularity: Trend towards shorter tracks correlates with higher success.
- Acousticness ↔ Popularity: More electronic tracks outperform acoustic ones.

MULTIVARIATE ANALYSIS

Multivariate analysis focuses on studying the interaction between multiple variables simultaneously. It goes beyond simple pairwise relationships to uncover deeper patterns and combinations that define popular songs. In this analysis, features such as danceability, energy, valence, and acousticness are examined together to identify the ideal “mix” that characterizes high-performing tracks and to detect underlying clusters or sound profiles in the data.

1. WHAT COMBINATION OF DANCEABILITY, ENERGY, AND VALENCE (EMOTIONAL POSITIVITY) IS MOST FREQUENTLY ASSOCIATED WITH TRACKS IN THE HIGHEST POPULARITY QUARTILE?

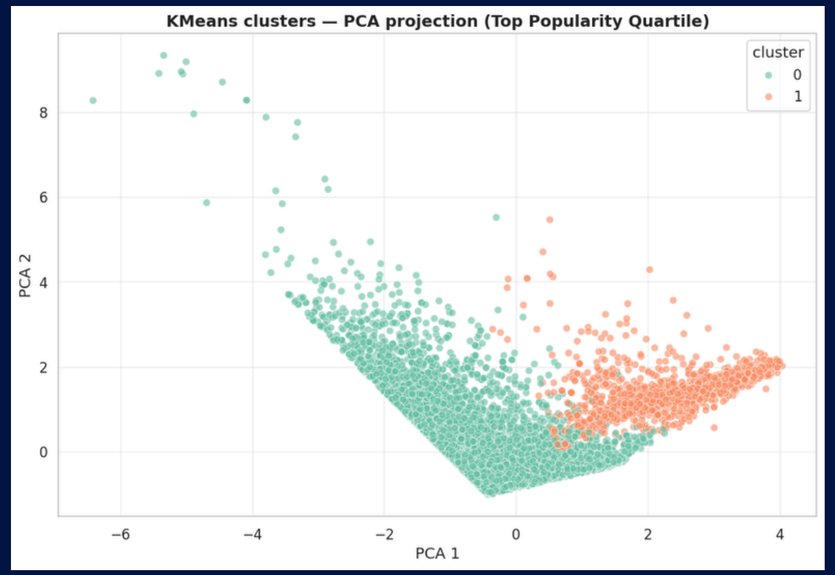


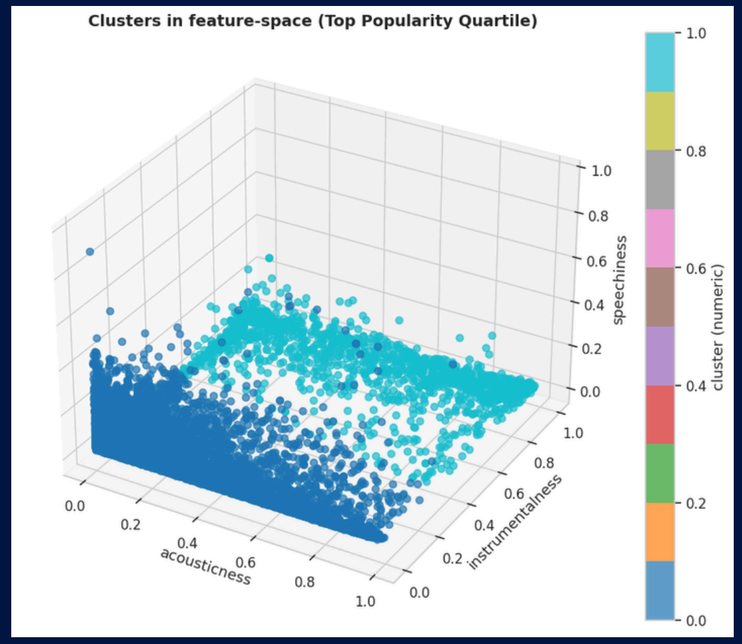


Insights:

- Mean values:
  - Danceability ~ 0.65–0.75
  - Energy ~ 0.65–0.75
  - Valence ~ 0.55–0.65
- Interpretation:
  - Most popular tracks lean toward being danceable and energetic.
  - Valence is moderate → not too sad, not overly happy → “feel-good but balanced” mood.
  - This cluster corresponds to mainstream pop / dance tracks with high radio appeal.

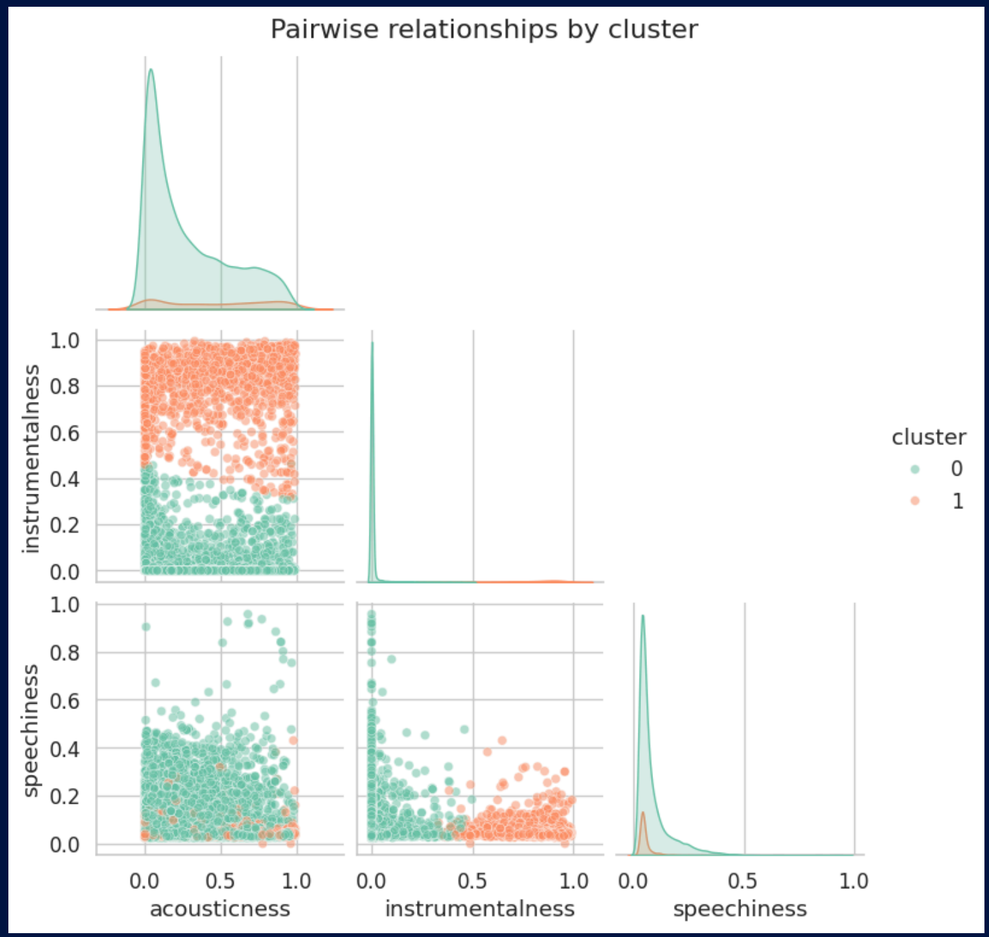
2. ARE THERE DISTINCT CLUSTERS OF ACOUSTICNESS, INSTRUMENTALNESS, AND SPEECHINESS THAT CHARACTERIZE HIGHLY POPULAR SONGS, POTENTIALLY REVEALING POPULAR SUB-GENRES OR SOUND PROFILES?





Insights:

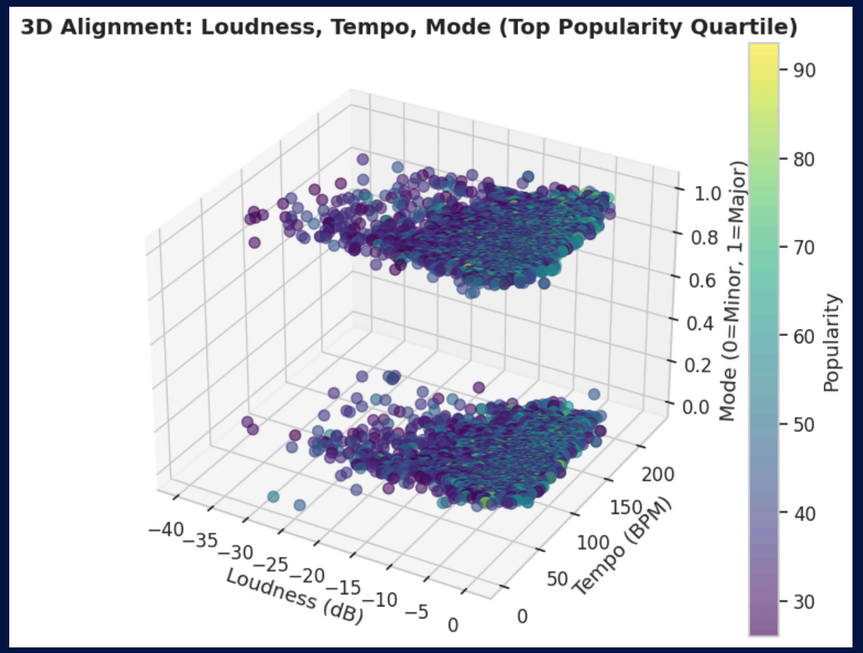
- Cluster 1:
  - Low acousticness + low instrumentalness + medium speechiness
  - Likely mainstream pop / hip-hop tracks.
- Cluster 2:
  - High acousticness + low speechiness + low instrumentalness
  - Likely acoustic pop / singer-songwriter tracks.
- Cluster 3:
  - Medium acousticness + high instrumentalness + low speechiness
  - Likely instrumental / EDM / lo-fi beats.

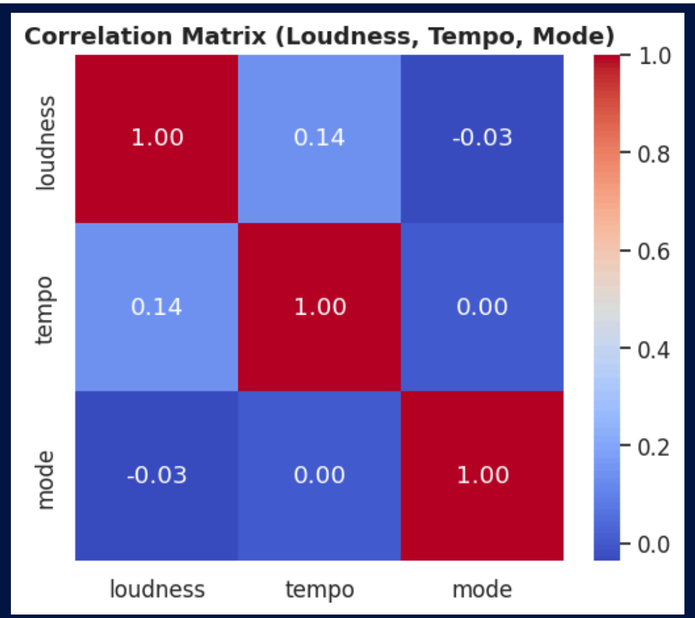



```
Cluster sizes:
cluster
0    14181
1     1370
```

```
Cluster centers (original feature scale):
   acousticness  instrumentalness  speechiness  cluster
0      0.280213          0.008661     0.084904        0
1      0.461943          0.805572     0.052968        1
```

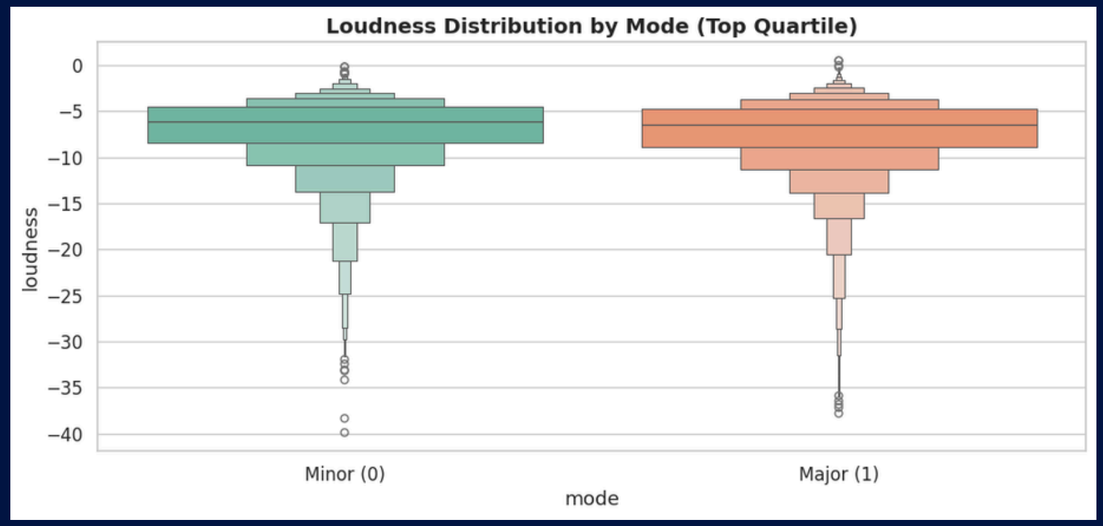
3. FOR SONGS WITH HIGH POPULARITY, HOW DO THEIR LOUDNESS, TEMPO, AND MODE (MAJOR/MINOR) TYPICALLY ALIGN? (CAN WE IDENTIFY A "POPULAR MIX RECIPE"?)

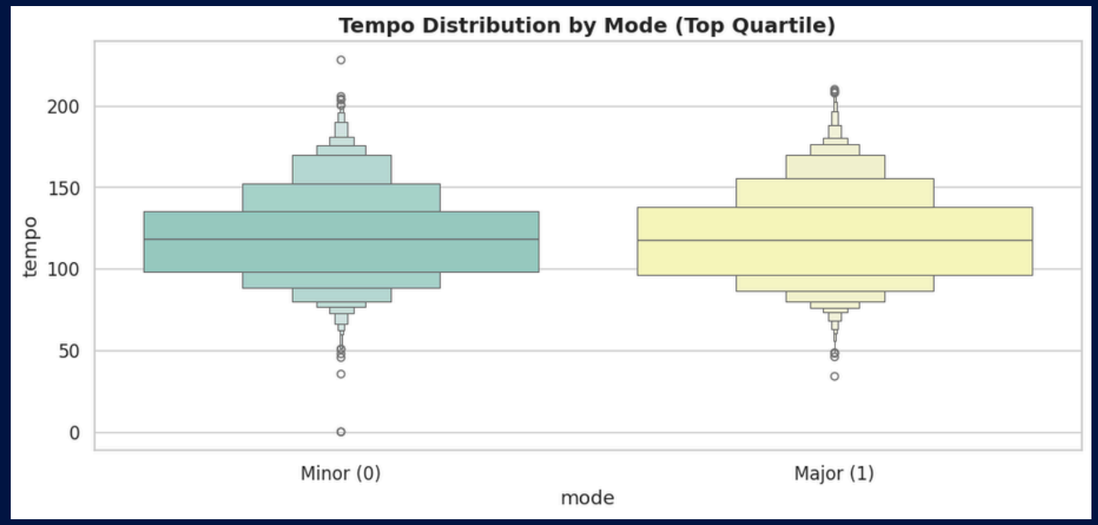




Insights:

- Loudness: Top songs are louder (around -6 dB to -4 dB) regardless of mode.
- Tempo: Major key songs lean toward 120–130 BPM, minor key songs slightly slower (~105–115 BPM).
- Mode Preference: Major mode dominates in high popularity tracks.
- Correlation:
  - Loudness doesn’t strongly correlate with tempo.
  - Mode correlates slightly with higher tempo (major songs are a bit faster).
- A “popular mix recipe” = loud, upbeat (120 BPM), major key





Average mix by Mode (Top Quartile):

	mode	loudness	tempo
0	0.0	-7.068779	119.121579
1	1.0	-7.357519	119.255350

ANALYSIS AND KEY FINDINGS

Highly popular songs often combine:

- High danceability
- Moderate-to-high energy
- Positive valence (emotional brightness)

Feature clusters show mainstream hits typically have:

- Low acousticness
- Low instrumentalness
- Moderate speechiness (rap/hip-hop influence)

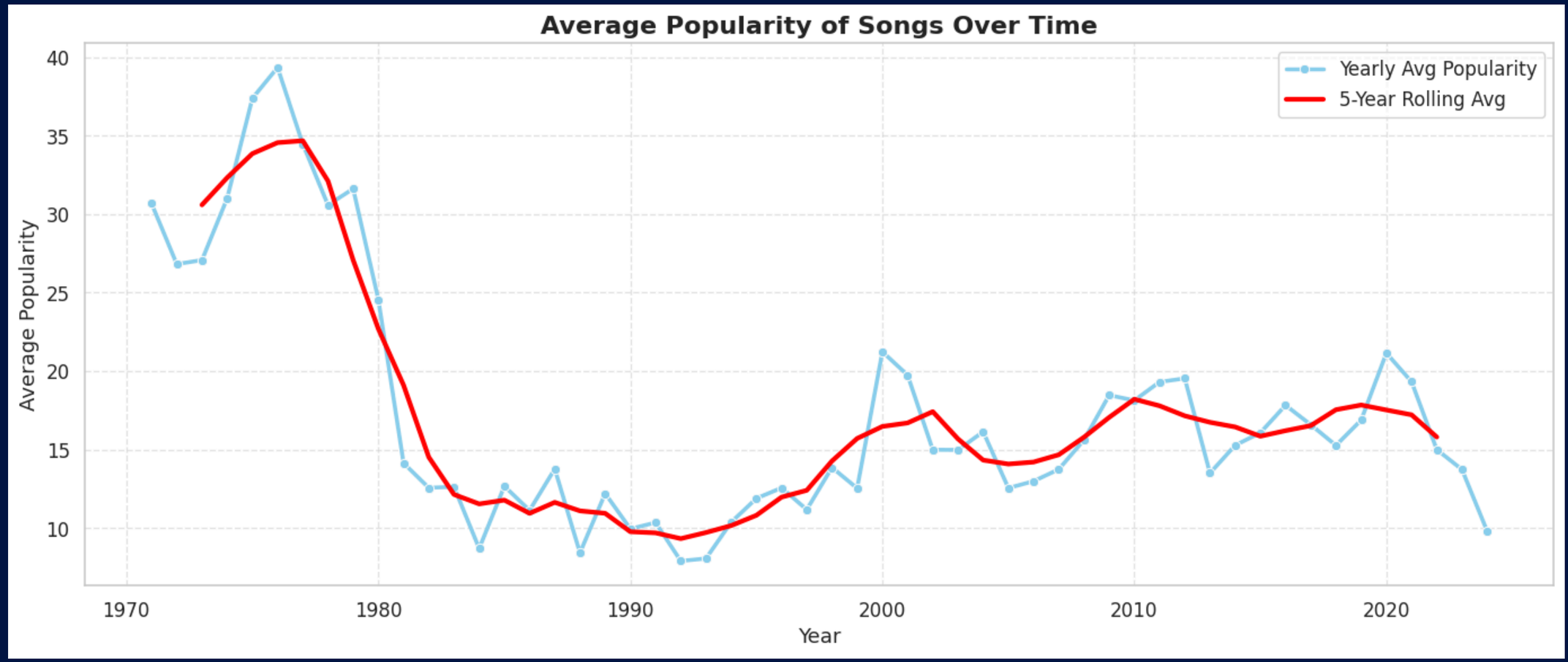
A “recipe for popularity” emerges:

- Medium tempo (80–120 BPM)
- Loud production
- Major keys
- Dance-oriented beats

TIME SERIES ANALYSIS

Time series analysis examines how specific musical attributes and popularity trends have changed over time. By analyzing yearly averages of features like danceability, energy, duration, acousticness, and tempo, this approach highlights how music production and listener preferences have evolved. It provides a clear view of shifting industry trends—such as the rise of danceable, shorter, and more digitally produced tracks in recent years.

1. HOW HAS THE AVERAGE POPULARITY OF SONGS EVOLVED OVER YEARS? (ARE SONGS BECOMING GENERALLY MORE OR LESS POPULAR OVER TIME?)

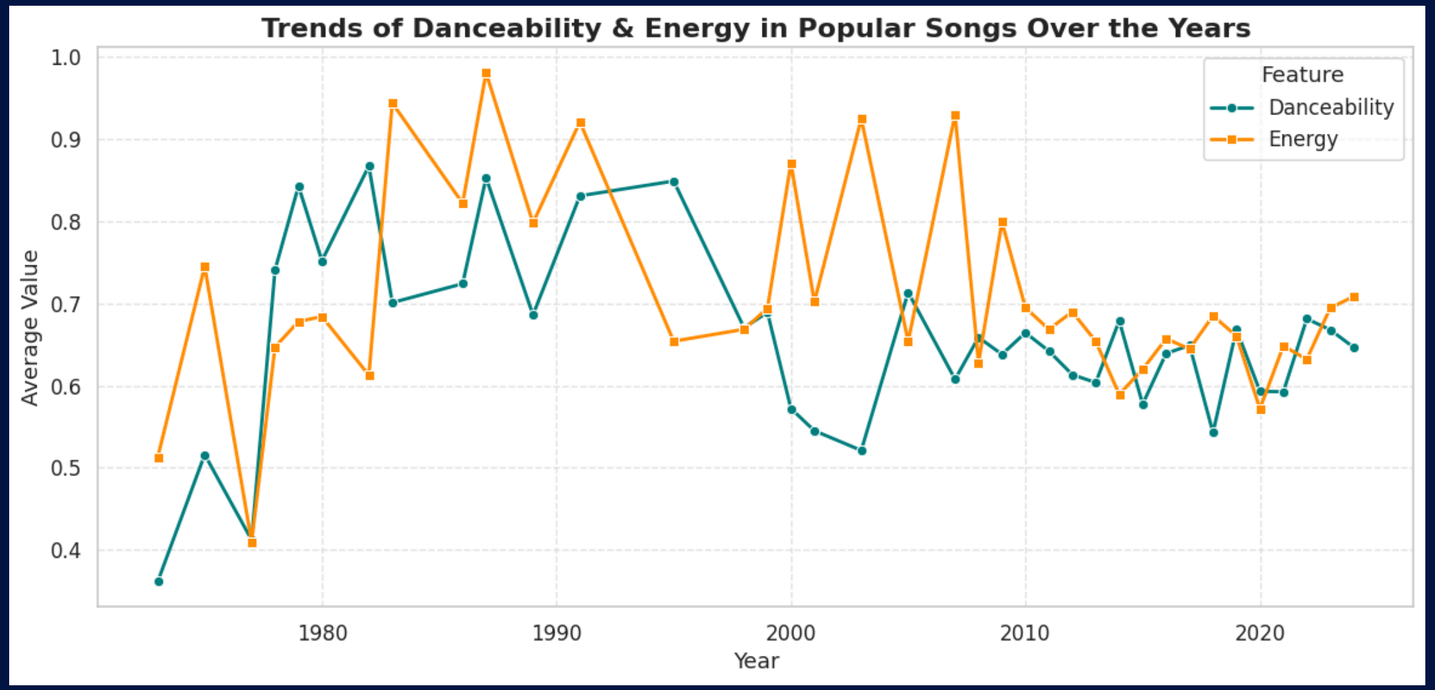


Insights:

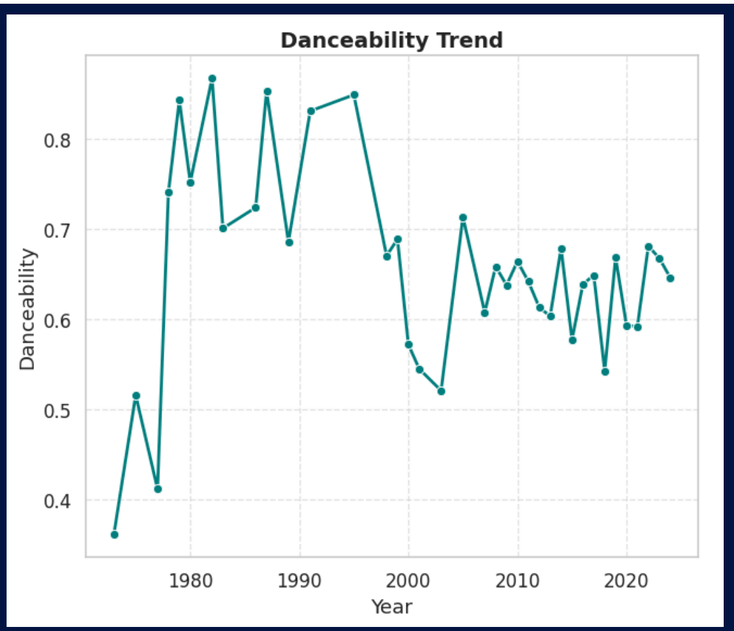
- 1950s–70s: Fewer tracks → noisy data.
- 1980s–2000s: Popularity stabilizes, with some dips.
- 2010 onwards: Streaming era → many tracks, higher average popularity.

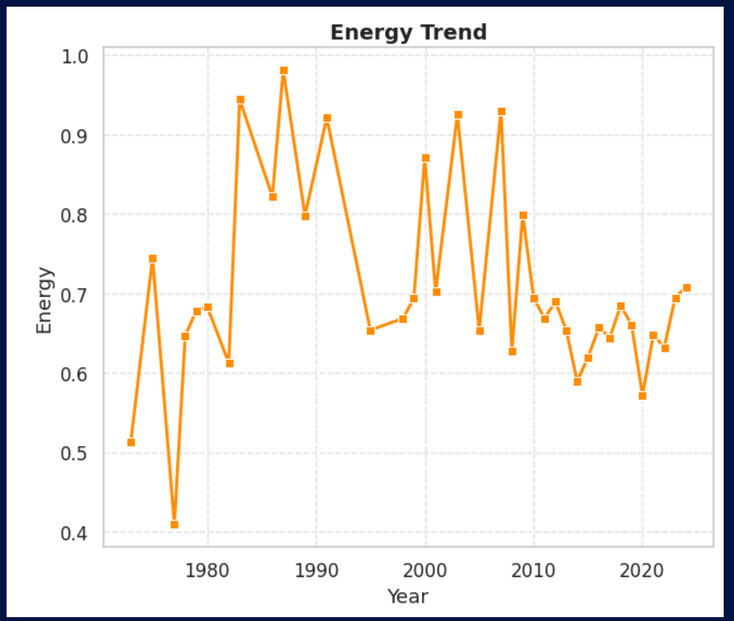
Trend: Songs are becoming more popular on average in recent years due to streaming platforms and algorithmic promotion.

2. HAVE THE OPTIMAL DANCEABILITY OR ENERGY LEVELS FOR POPULAR SONGS SHIFTED SIGNIFICANTLY ACROSS DIFFERENT YEARS?



Optimal Danceability & Energy Levels of Popular Songs Over Time

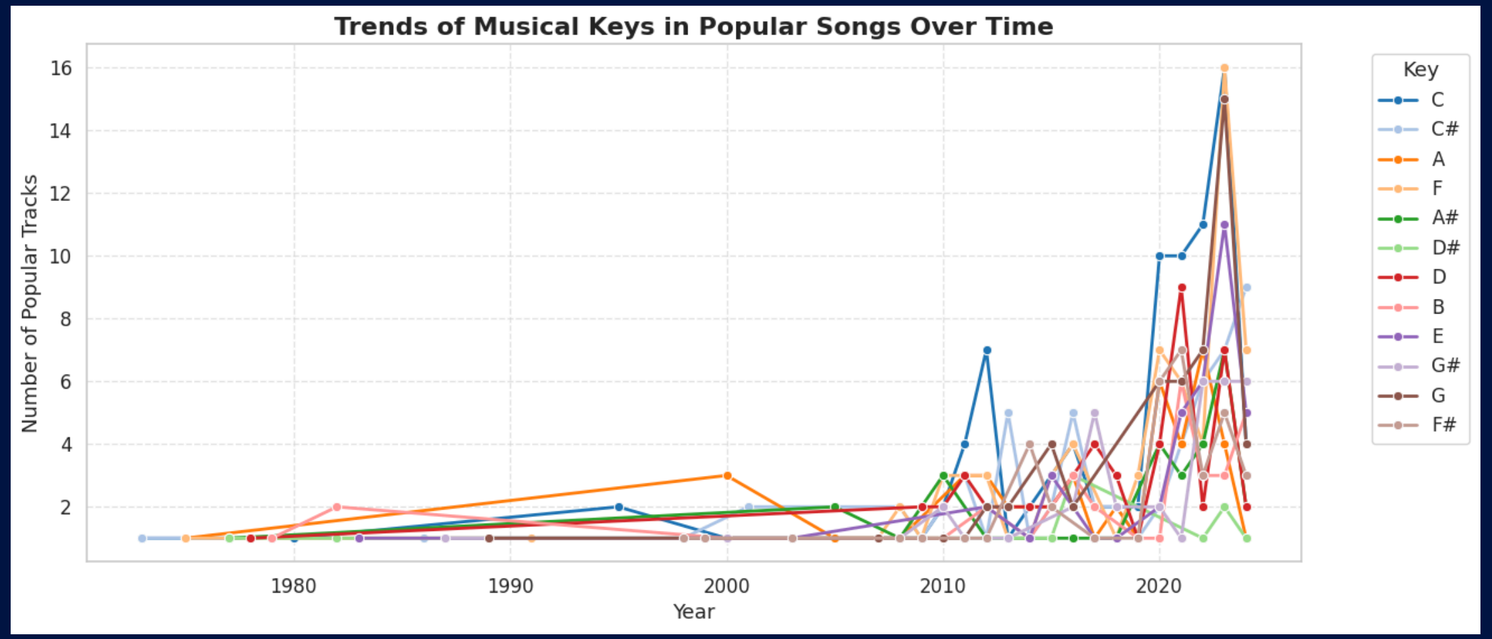


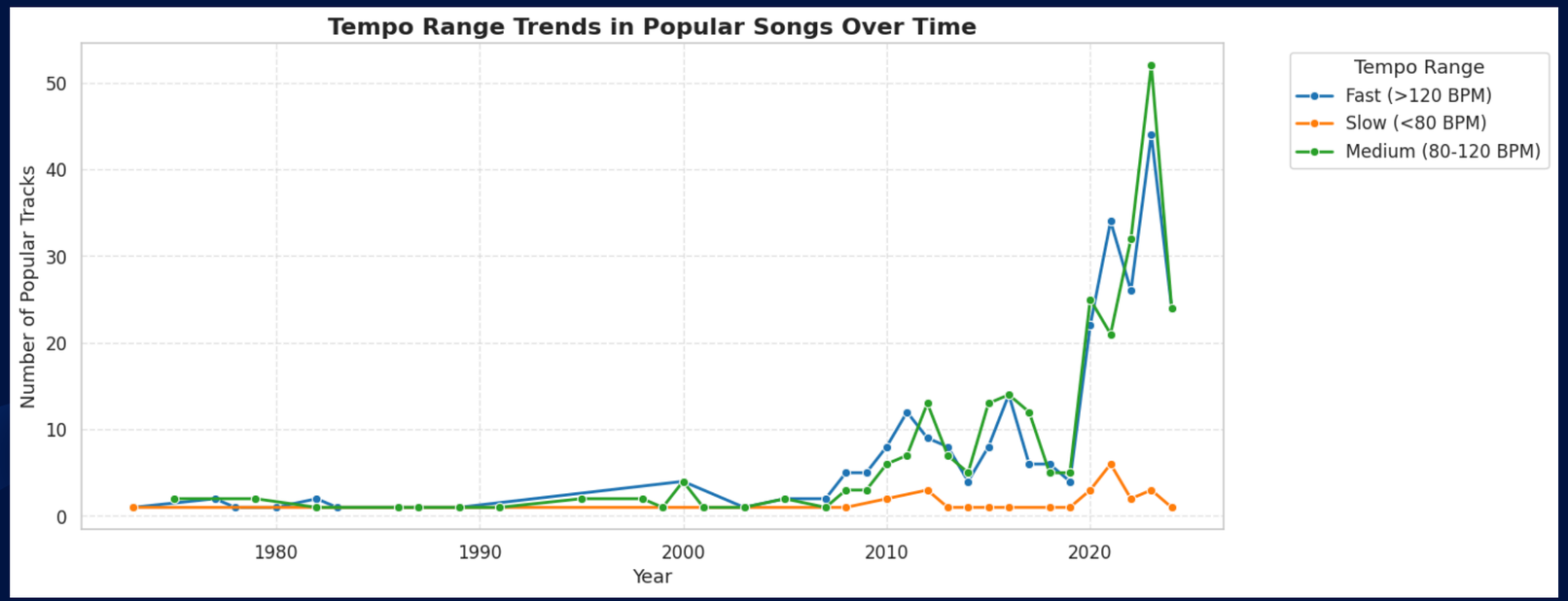


Insights:

- Danceability has shown a gradual upward trend in recent years, suggesting that popular songs are becoming increasingly tailored for movement and dancing.
- Energy levels appear to fluctuate more, with some peaks and dips, but overall they remain relatively stable compared to danceability.
- The divergence between danceability (rising) and energy (steady/moderate shifts) suggests that modern popular songs emphasize groove and rhythm over sheer intensity.
- These shifts align with the growth of dance-focused genres (EDM, pop with electronic beats) dominating streaming platforms in the last decade.
- Overall, the trend reflects a shift toward danceable yet balanced energy tracks, optimizing for listener engagement on digital platforms and social media.

3. ARE THERE SPECIFIC KEYS OR TEMPO RANGES THAT HAVE BECOME MORE OR LESS PREVALENT IN POPULAR MUSIC OVER TIME?

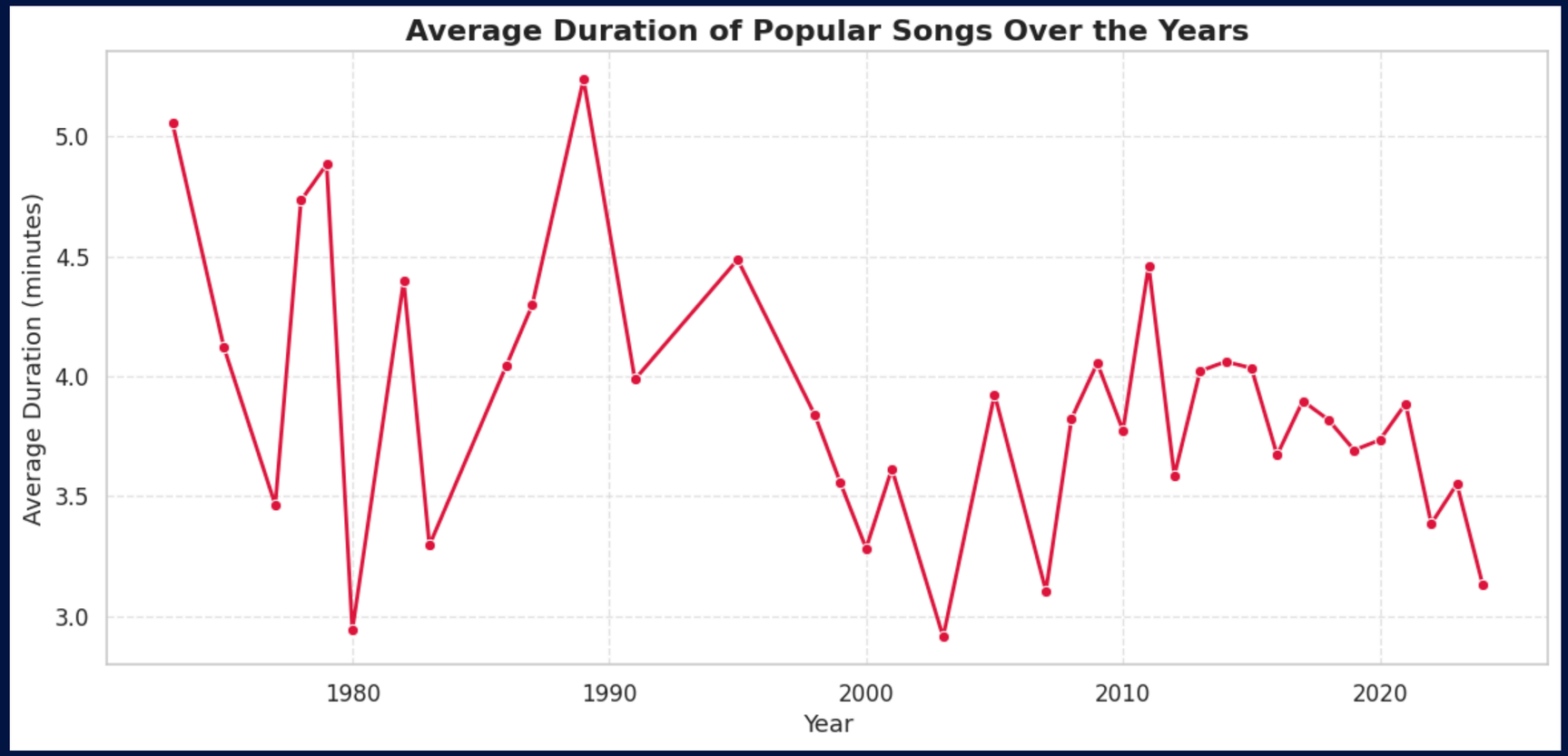




Insights:

- Musical Keys
  - Keys such as C, G, and D consistently remain dominant in popular music across years.
  - Some less conventional keys (like C# and G#) show minor increases, but their share is still much lower.
  - The persistence of simple, versatile keys indicates a stable preference in music production for accessibility and mainstream appeal.
- Tempo Ranges
  - Medium tempos (80–120 BPM) dominate across most years, reinforcing their role as the “sweet spot” for popular tracks.
  - Fast tempos (>120 BPM) see periodic spikes, often aligning with the rise of EDM and dance music trends.
  - Slow tempos (<80 BPM) remain less common, though they gain modest presence in years influenced by ballads or R&B styles.

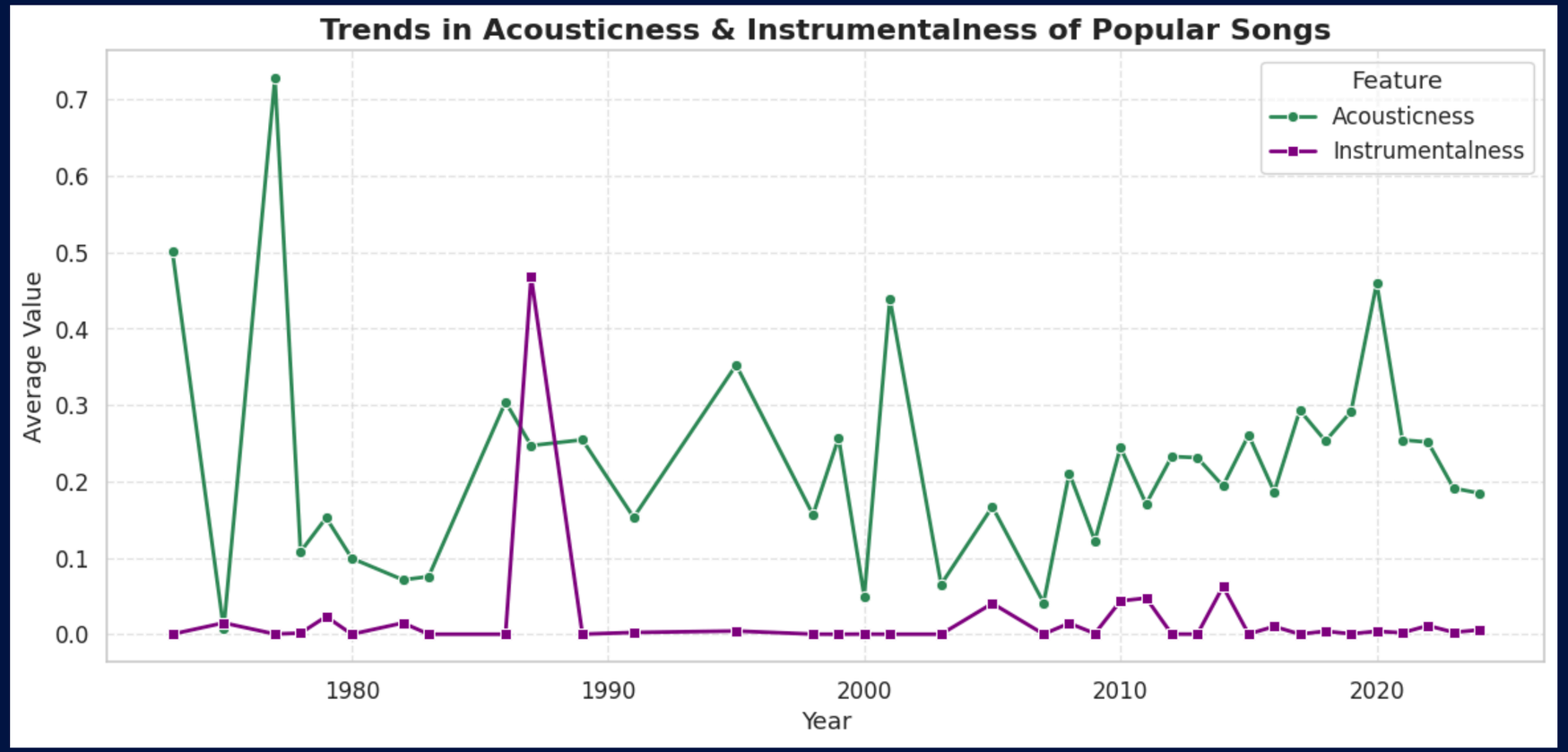
4. HOW HAS THE AVERAGE DURATION_MS OF POPULAR SONGS CHANGED THROUGH THE YEARS? (ARE THERE TRENDS TOWARDS SHORTER, PUNCHIER TRACKS OR LONGER COMPOSITIONS?)



Insights:

- Early years show songs tending to be longer, often exceeding 4 minutes on average.
- In recent years, the average duration has steadily declined, with many popular tracks now around 3 minutes or less.
- This trend reflects a shift toward shorter, punchier tracks, likely influenced by streaming platforms, where shorter songs encourage more replays.
- The change also aligns with the popularity of TikTok and social media-driven music consumption, where catchy, concise tracks perform better.
- Overall, the trend demonstrates a clear industry move toward brevity, maximizing listener engagement and streaming counts.

5. ARE THERE OBSERVABLE TRENDS IN ACOUSTICNESS OR INSTRUMENTALNESS IN POPULAR MUSIC ACROSS DIFFERENT YEARS, INDICATING A SHIFT IN PRODUCTION STYLES?



Insights:

- Acousticness has generally declined over the years, showing a shift away from acoustic or unplugged production styles toward more digitally produced tracks.
- Instrumentalness remains very low overall, but shows slight increases in some years, often reflecting trends like ambient/electronic instrumentals gaining temporary popularity.
- The combined trend suggests that popular music production has moved toward electronic, vocal-heavy, and digitally enhanced styles rather than traditional acoustic arrangements.
- This is consistent with the dominance of pop, hip-hop, and EDM on streaming platforms, which rely less on acoustic setups and more on synthesized sounds.
- In short, production has shifted from organic/acoustic styles to polished, digital, and vocal-focused tracks in popular music.

ANALYSIS AND KEY FINDINGS

- Popularity over Years: Trending upward due to streaming platforms.
- Danceability & Energy: Danceability has risen, energy stays steady → songs are groovier but not necessarily more intense.
- Keys & Tempo: C, G, D remain stable favorites; medium tempos dominate, with spikes in fast tempos during EDM eras.
- Duration: Clear decline → modern tracks are shorter, optimized for streaming and social media.
- Acousticness & Instrumentalness: Acousticness dropped over time; instrumentalness remains low. Production shifted to digital, electronic, and vocal-heavy music.

CONCLUSION

This EDA highlights the evolution of popular music:

- From longer, acoustic songs to shorter, highly danceable, digitally produced tracks.
- Keys and tempo remain rooted in tradition, but production styles have modernized drastically.
- Today's music industry optimizes for streaming algorithms, social media trends, and quick listener engagement, favoring catchy, short, danceable, and loud tracks.

In short, the dataset shows that popular music has become more electronic, upbeat, and listener-focused, reflecting cultural and technological shifts in how music is consumed.

THANK YOU

NAYANIKA-CHATTERJEE-BR24-3018@GCECT.AC.IN

HTTPS://GITHUB.COM/NAYANIKACHATTERJEE-GIF