

## Data Source:

[Top Hits Spotify from 2000-2019](#)

[Billboard "The Hot 100" Songs | Kaggle](#)

## Attributes:

### Spotify:

1. **artist**: Name of the Artist.
2. **song**: Name of the Track.
3. **duration\_ms**: Duration of the track in milliseconds.
4. **explicit**: The lyrics or content of a song or a music video contain one or more of the criteria which could be considered offensive or unsuitable for children.
5. **year**: Release Year of the track.
6. **popularity**: The higher the value the more popular the song is.
7. **danceability**: Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.
8. **energy**: Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity.
9. **key**: The key the track is in. Integers map to pitches using standard Pitch Class notation. E.g. 0 = C, 1 = C#/D♭, 2 = D, and so on. If no key was detected, the value is -1.
10. **loudness**: The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typically range between -60 and 0 db.
11. **mode**: Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.
12. **speechiness**: Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks.
13. **acousticness**: A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.
14. **instrumentalness**: Predicts whether a track contains no vocals. "Ooh" and "aah" sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly "vocal". The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0.

15. **liveness**: Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live.
16. **valence**: A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).
17. **tempo**: The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.
18. **genre**: Genre of the track.

#### **BillBoard:**

1. **Track Name**: The name of the song.
2. **Artist Name**: The name of the artist or group.
3. **Last Week Position**: The position of the song in the Billboard Hot 100 chart during last week.
4. **Peak Position**: The highest rank that the song achieved on the chart.
5. **Year**: The year the song was charted.

### Why This Dataset:

#### **Spotify Dataset:**

- Provides insights into the evolution of popular music on Spotify.
- Allows for the analysis of trends in various music characteristics over time, such as **danceability**, **energy**, and **tempo**.
- Enables the study of how attributes like **loudness** and **danceability** correlate with a song's popularity.
- Helps explore the changing preferences of listeners, including shifts in **musical styles** and individual song characteristics.

#### **Billboard Dataset:**

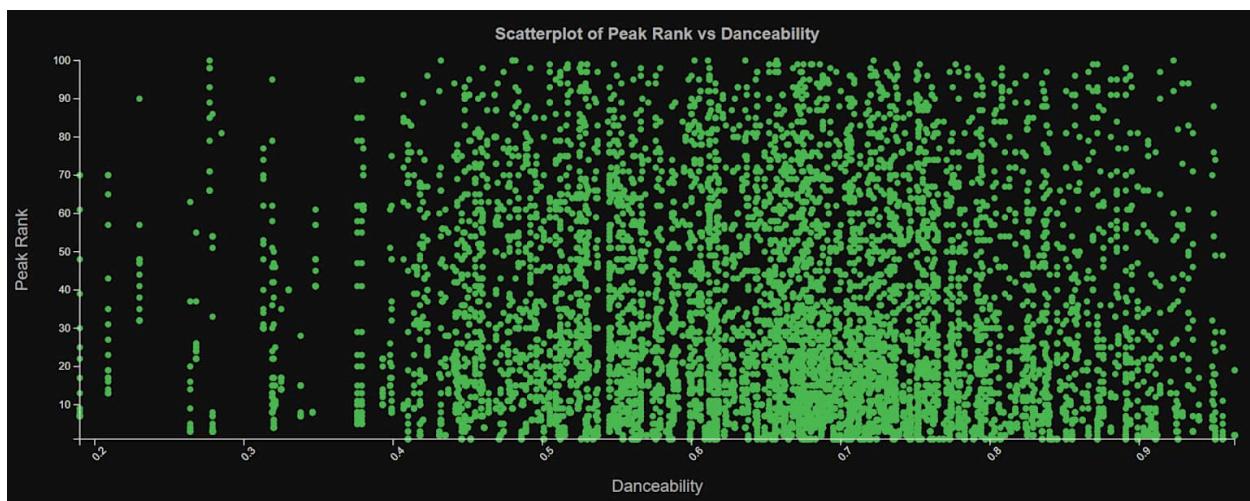
- Offers a historical perspective on the songs that have been popular in the U.S. over time.
- Provides a deeper exploration of music trends, focusing on how songs and artists performed in different years and the **longevity** of a song's success.
- Analyzing **weeks on the chart** and **peak position** offers insights into what factors contribute to a song's staying power or breakthrough performance.
- **Genre** and **record label** data allow for analysis of how different music types and industry support influence chart performance.

#### **Benefits of Combining Both Datasets:**

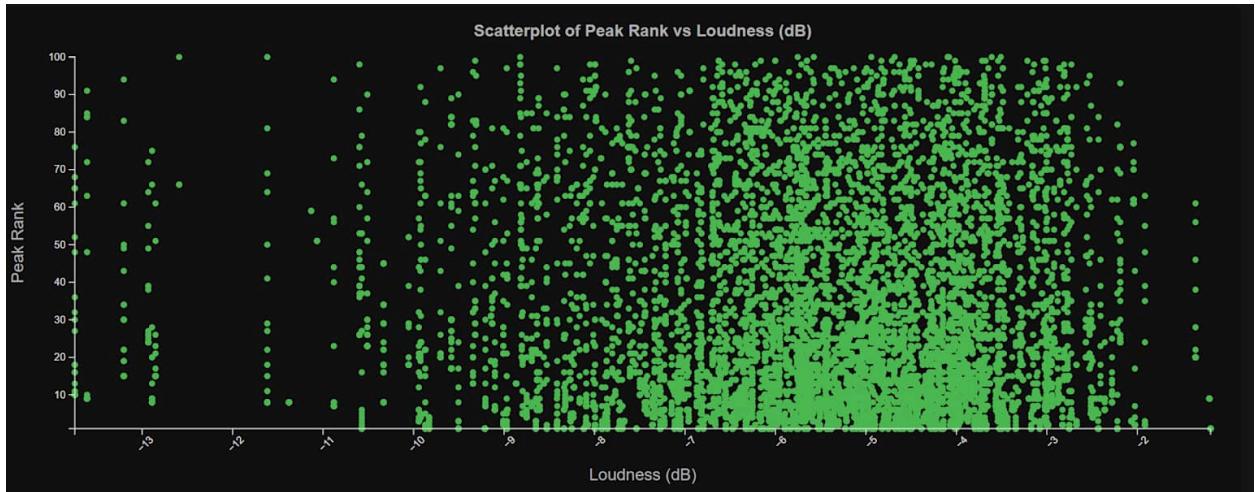
- **Cross-Platform Comparison:** Allows for comparing how songs perform on both **Spotify** and **Billboard** charts, providing a clearer picture of their success across different platforms.
- **Deeper Analysis of Success Factors:** Facilitates an analysis of **longevity** and factors that contribute to a song's success, such as **genre**, **popularity trends**, and musical characteristics like **danceability**, **energy**, and **tempo**.
- **Predictive Analysis:** Combining chart performance and streaming success data allows for predicting which songs will likely thrive on both platforms and the role of **marketing**, **artist**, and **label** strategies in influencing cross-platform success.
- **Comprehensive View of Music Evolution:** Offers a broad perspective on how music trends have evolved, particularly in relation to the growing influence of **streaming services** like Spotify.
- **Enriched Metrics:** Provides more detailed metrics for better decision-making

## Noteworthy Observation:

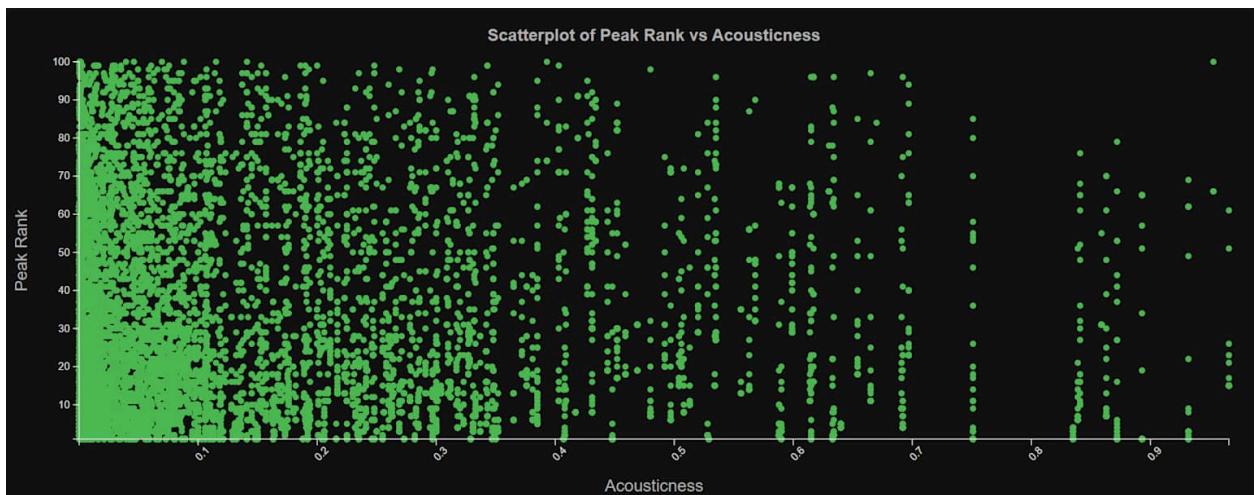
Let us look into some of the plots we derived with our dashboard:



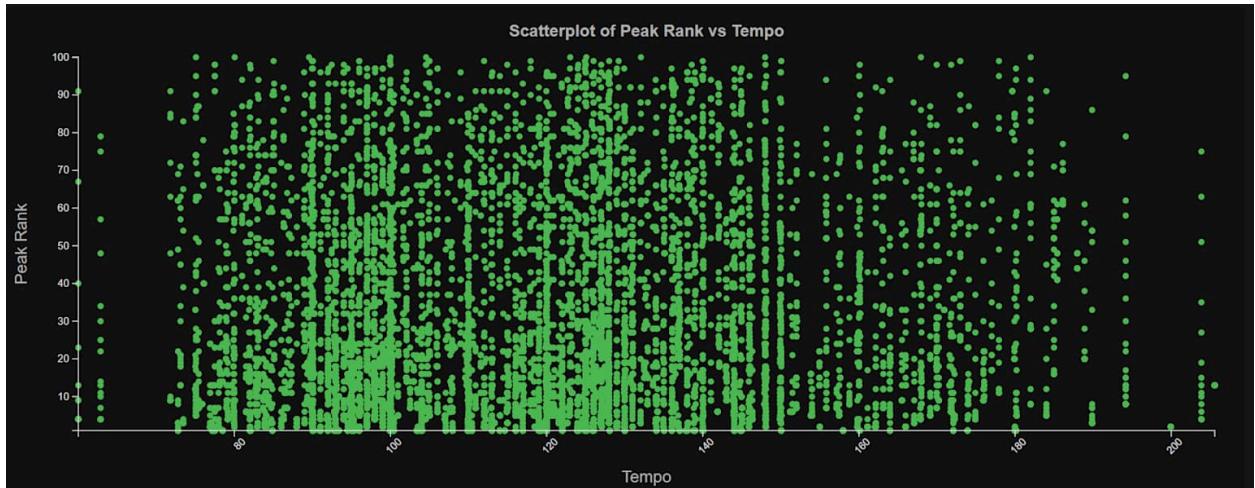
- There's a strong positive correlation between danceability and peak rank performance
- Songs with danceability scores above 0.6 tend to perform better on the charts
- The highest concentration of successful songs (high peak ranks) occurs in the 0.6-0.8 danceability range
- Very low danceability scores (below 0.4) rarely result in high-charting songs



- Most successful songs cluster in the -7 to -3 dB range
- There's a moderate positive correlation between loudness and chart performance
- Extremely quiet songs (below -12 dB) rarely achieve high chart positions
- The data suggests an optimal loudness "sweet spot" for commercial success



- There's a notable negative correlation between acousticness and peak rank
- Most highly successful songs have low acousticness scores (below 0.3)
- Very acoustic songs (scores above 0.8) rarely reach high chart positions
- This suggests modern popular music tends to favor electronic-produced sounds over acoustic ones



- The relationship between tempo and chart success is more complex and less linear
- Most successful songs cluster between 90-140 BPM
- There are distinct "columns" in the data, suggesting certain specific tempos are particularly common
- Both very slow (below 80 BPM) and very fast (above 160 BPM) songs are less common in high chart positions