

Project Proposal

on

Billboard Through the Ages: Charting the Evolution of Music Trends

By

Team:

Ayesha Tajammul Ahmed Mulla (amulla@iu.edu)
Shreya Gajbhiye (sygajbhi@iu.edu)
Saradhi Somarouthu (ssomaro@iu.edu)

Abstract

Intrigued by the generational divide in musical tastes, this project delves into the perceived shift in music quality and content over the years. We aim to explore the intricate dynamics of popular music spanning two decades, from 1999 to 2019. Older generations often find modern music too loud and lacking substance while expressing nostalgia for the meaningful lyrics and soulful artists.

By using data from Billboard's Hot 100 charts, Grammy Song Awards and Spotify's API, for analyzing the interplay between qualitative and quantitative aspects of music, we want to understand how music has changed over the last 2 decades. Through rigorous data analysis and visualization techniques, we intend to uncover long-term trends, patterns, and shifts in music popularity. Our work focuses on understanding the factors that contribute to a song's trendiness, popularity on Billboard, and its recognition through prestigious awards like Grammys.

To achieve these objectives, we will employ a series of data visualization methods. These visualizations will illuminate the connections between song attributes, genres, artists, and chart performance. Moreover, we extend our study to investigate multiple facets of music and study its impact on a song's popularity, that connects with listeners, ensuring songs remain on charts and gain critical recognition. By combining the qualitative insights from Billboard and Grammy data with the quantitative attributes provided by Spotify, our project offers a holistic view of the music industry's evolution to music enthusiasts.

Introduction

In our daily lives, music permeates every aspect, providing the rhythm to our existence. Music is more than just a collection of sounds; it is a reflection of culture, a symbol of change, and a representation of individual and collective preferences. They are cultural landmarks that define eras and often serve as vehicles for societal shifts.

This project embarks on a journey to unravel the perceived shift in music over the course of two decades, spanning from 1999 to 2019, a time when pop sensations and rock blockbusters dominated the radio waves. This ever-changing soundscape was retained by the Billboard Hot 100 rankings, which served as a monthly record of the biggest melodies, while the Grammy Awards honored musical icons, thus immortalizing the names of legends in music history. What unfolds when we bridge these two domains, where chart-topping rhythms meet award-winning acclaim?

Consider the story of Beyoncé, Taylor Swift, and Adele, all musical powerhouses. They each have their own unique Grammy stories. Beyoncé has a whopping 32 Grammys, but surprisingly, only one of them falls into the "Big Four" categories. Adele, on the other hand, boasts seven such wins, surpassing both Swift and Beyoncé. Swift, though, stands out as the only woman to clinch the album of the year title three times, a feat shared only with legendary male artists.

This intriguing interplay between Billboard chart-toppers and Grammy victors raises questions about the changing landscape of music over time.

What makes a song popular? Why are certain songs on the Billboard charts? Why do certain artists win Grammy Awards? To understand this, we should dive deep into the intricacies of a piece of musical art.

Each song is like a unique puzzle, and the pieces of that puzzle are what we refer to as "song attributes." These attributes serve as special tools that help us decipher and appreciate music on a deeper level.

Think of "Beats Per Minute" or BPM as a song's speed, just like how fast a car is going. "Energy" tells us how lively a song feels, with higher values meaning more excitement – like when your favorite team scores a goal. "Danceability" is like an invitation to dance; some songs make us dance easily, while others don't. "Loudness" measures how loud or quiet a song is, just like adjusting the volume on your TV. "Valence" gives us clues about a song's emotional tone, helping us figure out if it's happy or sad. Then there's "Length," which is simply how long the song is. And "Acousticness" shows us if a song is more acoustic or electronic. Lastly, "Release Year" gives us a peek into when the song was made. These tools help us see music in a new way, like reading a book and learning about its author, genre, and the time it was written.

When we put all these song attributes together, they reveal fascinating stories about music. These stories take us through how music has changed over time and help us understand why some songs make us want to dance while others make us feel calm and reflective. These song attributes are like a secret code, letting us unlock the enchanting world of music and discover the magic behind every musical note.

In this project, as we create charts and graphs, we reveal the rise and the fall of genres, the beats that make us dance, and the emotions that resonate with our hearts. These visual cues not only entertain but also educate us by offering insights into why some tunes top the charts while others fade into the background. By depicting music through visualizations, we are trying to make these musical narratives accessible to all, leading to discovery.

Visualization Critiques:

Critique 1:

The visualization "Which Month Has the Most New Music Releases?" displays trends in music release patterns across different months, covering several decades. It uses a line graph where each month is represented by a small colored circle, showing the variations in music releases over the years. However, it is challenging to interpret and understand the trends accurately because of the following reasons:

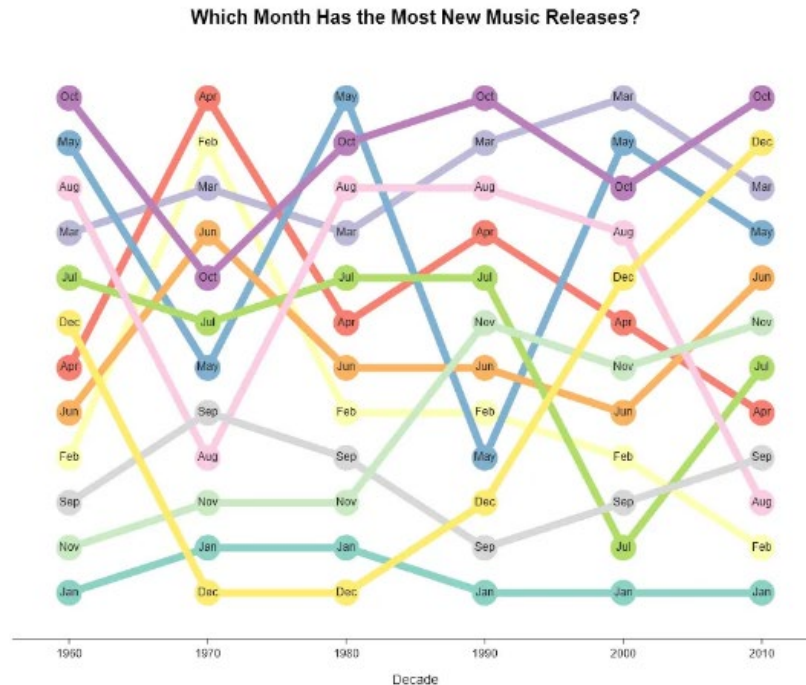


Figure: Line Graph for “Music Release Pattern Across Different Months”

1. **Unlabeled Y-Axis:** The Y-axis is not labeled at all, leaving viewers without a clear understanding of what the vertical axis represents. This lack of labeling makes it challenging to interpret the data accurately.
2. **Lack of a Clear Legend:** There is no legend or key to explain the color-coding used for each month. This leads to confusion regarding the significance of each color.
3. **Inadequate Color Scheme:** The color scheme employed in the visualization is problematic. While colors are used to distinguish between months, they may not be easily distinguishable for color-blind individuals. Using this color scheme as the primary encoding method limits accessibility and inclusivity.
4. **Limited Readability:** The visualization uses small, circular data points to represent each month. These points can be challenging to discern, especially when several overlap. This leads to a lack of readability and difficulty in extracting meaningful insights.

The stacked bar chart would help to improve the existing visualization where each month can be a separate bar with segments for different years. This approach will enhance accessibility while providing clear labels for months and years, making it easier to comprehend.

Critique 2:

The data visualization titled “Genres of bands in each decade” is a Stacked Bar Chart representing popularity of bands belonging to various genres in every decade. Before the 90s, approximately 60% of bands were rock bands that typically consisted of one lead singer and a bunch of instrumentalists. With the progression of time in 2000s the percentage of rock bands reduced significantly leading to the rise of new

evolving brands of bands especially Pop. This indicates that the use of instruments dropped as roc bands became less popular with time. This visualization is good and useful for following reasons:

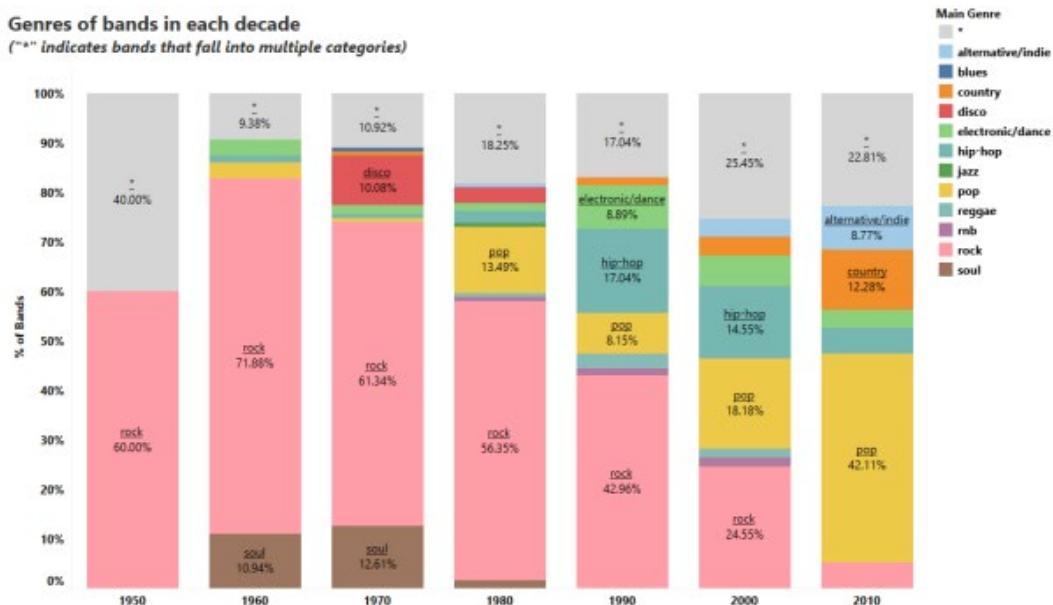


Figure: Stacked bar Chart for “Genres of bands in each decade”

1. **Effective Use of Color:** Using different colors for each genre helps in distinguishing between categories. This color coding enhances readability and makes it visually appealing.
2. **Percentage Labels:** Including percentage values on the bars provides precise information about the popularity of each genre. This data labeling enhances the accuracy of interpretation.
3. **Comparative Analysis:** It facilitates easy comparison between genres within the same year. Viewers can quickly identify which genres dominated in a specific period and how their popularity fluctuated.
4. **Clear Representation of Data:** The choice of graph is extremely suitable for representing this data. The stacked bar chart provides a clear visual representation of the popularity of bands from various genres over different years. Each bar is divided into segments, making it easy to see the composition of genres in a specific year.

The stacked bar chart visualization is helpful for our project as it guides us in data exploration phase to perform in-depth analysis. By combining these insights with our dataset, we can uncover intricate patterns, correlations, and historical contexts, enriching our project's narrative and offering valuable insights into the dynamic world of music trends.

Critique 3:

The data visualization titled 'The Rise and Fall of Genres' offers a comprehensive analysis of the Billboard top 20 chart, explaining the temporal trends observed within various music genres. This visualization is particularly good because of several key attributes:

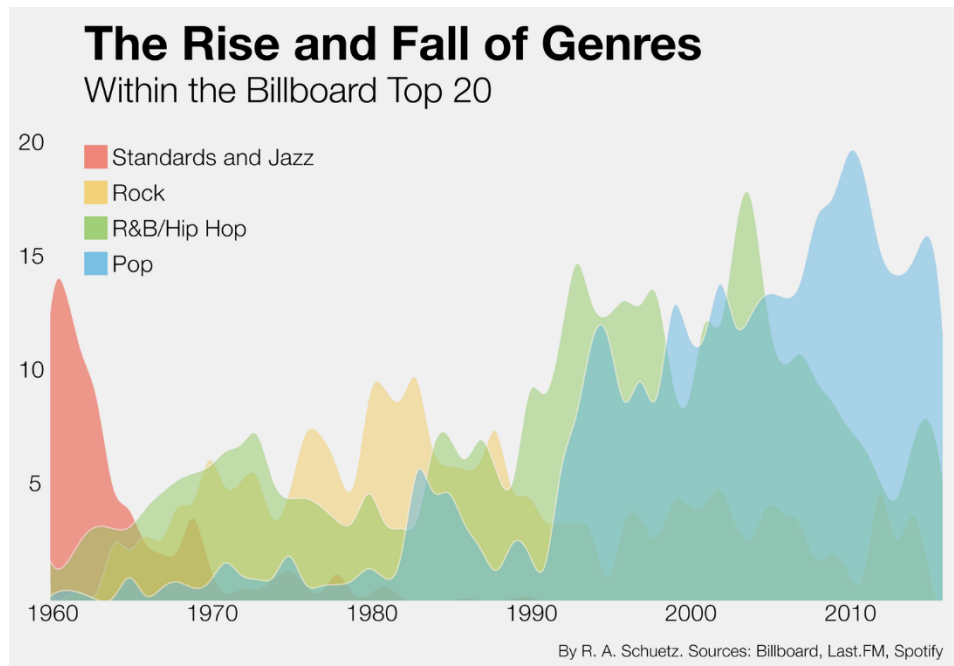


Figure: The Rise and Fall of Genres

1. **Simplicity and Intuitiveness:** The visual representation is notably straightforward, facilitating easy comprehension for the audience.
2. **Effective Utilization of Color Hues:** The visualization employs a sophisticated palette of color hues, skillfully leveraging the overlap of these hues to convey intricate data patterns.
3. **Minimalistic Design:** The presentation adopts a minimalist design philosophy, focusing on the essential elements while avoiding unnecessary clutter, resulting in a clean and uncluttered display.
4. **Rich Informational Content:** The visualization offers a wealth of information concerning listener preferences and trends, allowing for in-depth analysis of evolving musical genre preferences over time."

Reviewing existing visualizations, we find a mix of effective and lacking elements. For instance, the visualization titled 'The Rise and Fall of Genres' impressively simplifies complex data, utilizes color hues effectively, maintains a minimalistic design, and offers rich informational content. However, other visualizations, such as 'Which Month Has the Newest Music Releases?' suffer from unlabeled axes, lack of clear legends, problematic color schemes, and limited readability.

Thus, data may be made simpler to comprehend by identifying axes, applying accessible color schemes, and adopting clear visual encoding methods. Additionally, providing context and interactive elements can engage viewers and strengthen their knowledge about the data.

Our project bridges the motivation of understanding evolving music trends with existing visualizations. We aim to build upon and contribute to knowledge by creating more accessible, informative, and inclusive visualizations that provide a deeper understanding of music's transformation over time.

Questions/objectives

1. Evolution of Music Genres Over Time

Question: How have the song attributes of popular music changed over the past two decades?

Visual: Heatmap or animated graph

Insight: Identify changes in the popularity of different music genres over the years, revealing evolving musical preferences..

2. Artist's Grammy Success vs. Billboard Rankings

Question: How do the song attributes of award-winning songs compare to those of popular songs on Billboard?

Visual: Scatterplot

Insight: Explore potential correlations between the number of Grammy awards won by artists and their average Billboard Hot 100 chart rankings.

3. Lyric Sentiment Analysis Over Time

Question: Are there any common themes or topics in the lyrics of award-winning songs or over time?

Visual: Line chart

Insight: Analyze the emotional sentiments of song lyrics over time, highlighting trends in lyrical themes and moods.

4. Acousticness and Danceability by Genre

Question: What factors such as Acousticness and Danceability vary across Genres

Visual: Grouped bar chart

Insight: Compare the average acousticness and danceability scores across different music genres to identify genre-specific characteristics. Track how attributes like energy, instrumentalness, and valence have evolved over time, shedding light on changing musical styles.

5. Word Cloud of Popular Writing Credits

Questions: How does the Lyrics of the songs attribute to the popularity?

Visual: Word cloud

Insight: Showcase prolific songwriters by visualizing the frequency of their contributions to popular music.

Other Questions:

- What are the key factors that contribute to a song's trendiness and popularity on Billboard?
- How has the popularity of different music genres changed over time?
- Which artists have had the most Grammy-nominated songs that also reached the top 10 on the Billboard Hot 100 chart?
- What are the most popular genres of music on Spotify among Grammy-winning artists?
- Which artists are most likely to win Grammy Awards based on their current popularity on Spotify and the Billboard Hot 100 chart?
- What songs are most likely to become hits based on their musical characteristics and the popularity of similar songs on Spotify and the Billboard Hot 100 chart?

Dataset

For the project, our dataset comprises of three csv data files. The first file includes details about songs that made it to the Billboard Hot 100 chart from July 1999 to July 2019. This file contains information such as genre tags, song lyrics, and writing credits. The second file contains information about Grammy Awards won by artists for their songs. The third file provides song attributes obtained from Spotify, including features like danceability, instrumentalness, energy, etc.

Billboard Hot 100 csv:

Billboard Hot 100 dataset contains information about all the artists and their songs that featured on the Billboard Top 100 chart between the years 1999 to 2019. This dataset contains information on Popular Songs. There are 11 features in this dataset. These features are described below along with their data types.

- **Artists (String):** Name of artist(s).
- **Name (String):** Name of the song.
- **Weekly.rank (Int):** Rank on Billboard Hot 100 in a specific week.
- **Peak.position (Float):** Highest position the song has reached on the Billboard Hot 100 chart.
- **Weeks.on.chart (Float):** Weeks the song has been on the Billboard Hot 100 to this point.
- **Week (String):** Number of current weeks the song has been on the Billboard Hot 100.
- **Date (Date):** Release date of song.
- **Genre (String):** Genre tags of the song.
- **Writing.Credits (String):** List of people who contributed to writing the song lyrics.
- **Lyrics (String):** Lyrics of the song.
- **Features (String):** Additional artists or contributors who are featured in the song, but not its main performers.

Artists	Name	Weekly.rank	Peak.position	Weeks.on.chart	Week	Date	Genre	Writing.Credits	Lyrics	Features
Drake	Money In The Grave	8	7.0	2.0	2019-07-06	June 15, 2019	Hop,Rap,Basketball,NBA,Canada	Hip-Asoteric, Ljay currie, Cydney christine, Rick ...	Money in the Grave InYeah, okayInLil CC on the...	Rick Ross
Chris Brown	No Guidance	9	9.0	3.0	2019-07-06	June 8, 2019	Alternative R&B,Hip-Hop,Rap,Pop,DMV,Canada,R&B	Velous, J louis, Vinylz, Michee patrick lebrun...	No Guidance Before I die I'm tryna fuck you,...	Drake
Post Malone	Wow.	10	2.0	27.0	2019-07-06	December 24, 2018	Memes,Hip-Hop,Trap,Rap	Frank dukes, Billy walsh, Louis bell, Post malone	Wow InSaid she tired of little money, need a b...	NaN
Lizzo	Truth Hurts	11	11.0	8.0	2019-07-06	September 19, 2017	Pop,Rap	Ricky reed, Jesse saint john, Steven cheung, L...	Truth Hurts Why're men great til they gotta ...	NaN
Post Malone, Swae Lee	Sunflower	12	1.0	36.0	2019-07-06	October 18, 2018	Rap,Disney,Hip-Hop,Marvel,Soundtrack,R&B,Pop	Carl rosen, Louis bell, Billy walsh, Carter la...	Sunflower Ayy, ayy, ayy, ayy OohInOoh, ooh, ...	NaN

Figure: Sample Billboard dataset of 5 rows

Grammy Songs csv:

This dataset contains information on the grammy awards received by artists for their songs between 1999 to 2019. This dataset contains information on award-winning songs. There are 5 features in the dataset. These features are described below along with their data types.

- **GrammyAward (String):** Grammy award title.
- **GrammyYear (int):** Year.
- **Genre (String):** Genre of grammy award given.
- **Name (String):** Name of song/record that won the award.
- **Artist (String):** Name of artist who received the award.

GrammyAward	GrammyYear	Genre	Name	Artist
Record Of The Year	2018	General	this is America	Childish Gambino
Song Of The Year	2018	General	this is America	Childish Gambino
Best Pop Solo Performance	2018	Pop	Joanne (where Do you Think You're Goin'?)	Lady Gaga
Best Pop Duo/Group Performance	2018	Pop	Shallow	Lady Gaga & Bradley Cooper
Best Dance Recording	2018	Dance/Electronic Music	Electricity	Silk City & Dua Lipa Featuring Diplo & Mark Ronson
Best Rock Performance	2018	Rock	when Bad Does Good	Chris Cornell
Best Metal Performance	2018	Rock	Electric Messiah	High on Fire
Best Rock Song	2018	Rock	Masseduction	St Vincent

Figure: Sample Grammy Songs dataset of 8 rows

Spotify Song Attributes csv

This dataset contains information on various song attributes like accousticness, energy, danceability, instrumentalness, liveness and many more. These features were obtained from Spotify for songs released between 1999 to 2019. This dataset contains information for all songs. There are 17 features in this dataset. These features are described below along with their data types.

- **Acousticness (Float):** Measure of how acoustic a track is, ranging from 0 to 1.
- **Album (String):** Name of album.
- **Artist (String):** Name of artist.
- **Danceability (Float):** Indicates how suitable a track is for dancing, ranging from 0 to 1.
- **Duration (Int):** The duration of the track in milliseconds.
- **Energy (Float):** Represents a perceptual measure of intensity and activity of the track, ranging from 0 to 1.
- **Explicit (Boolean):** Whether the track is explicit.
- **Instrumentalness (Float):** Predicts whether a track contains no vocals, ranging from 0 to 1.
- **Liveness (Float):** Detects the presence of live audience in the track, ranging from 0 to 1.
- **Loudness (Float):** Represent the overall loudness of a track in decibels (dB).
- **Mode (Int):** 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.
- **Name (String):** Name of track.
- **Popularity (Int):** Spotify's rating of the track popularity as of July 2019.
- **Speechiness (Float):** Measures the presence of spoken words in the track, ranging from 0 to 1.
- **Tempo (Float):** The overall estimated tempo of a track in beats per minute (BPM).
- **TimeSignature (Int):** Indicates the number of beats in a bar, or a measure.

- **Valence (Float):** A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track.

Acousticness	Album	Artist	Danceability	Duration	Energy	Explicit	Instrumentalness	Liveness	Loudness	Mode	Name	Popularity	Speechiness
0.000728	Collective Soul (Deluxe Version)	collective soul	0.520	234947	0.904	False	0.010300	0.0634	-5.030	1	welcome all again	35	0.0309
0.018200	Collective Soul (Deluxe Version)	collective soul	0.581	239573	0.709	False	0.000664	0.1740	-4.909	1	fuzzy	31	0.0282
0.000473	Collective Soul (Deluxe Version)	collective soul	0.572	198400	0.918	False	0.000431	0.0977	-3.324	0	dig	30	0.0559
0.000970	Collective Soul (Deluxe Version)	collective soul	0.596	231453	0.661	False	0.000033	0.1130	-5.051	1	you	35	0.0254
0.000036	Collective Soul (Deluxe Version)	collective soul	0.520	222520	0.808	False	0.000010	0.0800	-4.553	0	my days	21	0.0318

Figure: Sample Song attributes dataset of 5 rows.

Kind of Visualizations that we plan to apply are as follows:

1. **Box Plot:**
 - Comparing song attributes like loudness, energy, etc. for various categories like genres.
 - Comparing song attributes for all songs, popular songs, and award-winning songs.
2. **Pie Chart:**
 - Display the distribution of award-winning songs won for different genres.
3. **Heatmap:**
 - Display the correlation between different song attributes.
4. **Scatter Plots:**
 - Explore the relationship between two continuous variables like danceability vs. popularity, energy of song vs. its popularity, peak position of songs vs. No. of weeks it was on Billboard chart, etc.
5. **Line Charts:**
 - Explore the current trends by plotting the relationship between the weekly rank of songs vs. the number of weeks it spent on Billboard chart.
 - Visualize the popularity of songs over years and display trends like weekly rankings on Billboard chart.
6. **Venn Diagram:**
 - Display the overlapping trends between award-winning songs and popular songs on Billboard by showing intersection between these songs through the Venn diagram.
 - Visualize the overlap of songs in different genres.
 - Visualize the overlap between lyricists and featured artists to explore which individuals wrote both lyrics and were featured in songs.
7. **Bar Charts and Histograms:**
 - Represent categorical data such as song genres featured on Billboard Hot 100 Chart using Bar Chart.
 - Explore the distribution of numerical variables such as song attributes using Histogram.
8. **Word Clouds**
 - Visualize the similarities and differences between artists who appear on Billboard and who get Grammy awards for their albums.

Why did we choose this Dataset?

This dataset is extremely well-suited for our project as it spans data over two decades between 1999 and 2019 offering a large amount of information allowing us to explore long-term trends, patterns, and shifts in music popularity, genres, and artist recognition. By leveraging data from Billboard and Grammy awards, we gain qualitative insights from the audience's perspective. Additionally, Spotify's song attributes dataset, offers ample information on quantitative aspects of songs. This dual perspective helps us achieve our objective to understand the impact of song attributes that contribute towards making the songs trendier and popular on Billboard leading to an award-winning recognition.

The usefulness of the dataset for this project is not only limited to the size of the dataset but also access to multitude of diverse features it offers to perform multi-faceted analysis, helping us in generating substantial evidence to support our observations.

Therefore, the dataset's depth and comprehensiveness empower us to learn and implement various data visualization techniques. Through these visualizations, we can unfold detailed and complex connections between song genres, chart the evolution of musical attributes, and investigate the intersection of how songs that are loved by listeners also gain critical recognition.

References

- <https://www.kaggle.com/danield2255/data-on-songs-from-billboard-19992019>
- <https://joshdviner.medium.com/what-makes-a-hit-song-analyzing-data-from-the-billboard-hot-100-chart-74c1d5ad3fa3>
- <https://www.billboard.com/lists/beyonce-taylor-adele-grammys-queen/most-overall-grammy-nominations/>
- <https://medium.com/cuepoint/visualizing-hundreds-of-my-favorite-songs-on-spotify-fe50c94b8af3>
- <https://towardsdatascience.com/hot-or-not-analyzing-60-years-of-billboard-hot-100-data-21e1a02cf304>
- <https://towardsdatascience.com/billboard-hot-100-analytics-using-data-to-understand-the-shift-in-popular-music-in-the-last-60-ac3919d39b49>
- <https://raschuetz.com/blog-fulton/2016/8/6/the-rise-and-fall-of-genres>