Temporal Trends in Popular Music Genres: A Data-Driven Analysis

Elizabeth Hall - Data Analysis Project

Elizabeth Hall

Invalid Date

# 1. Introduction

## 1.1 Background

The landscape of popular music has undergone significant transformations since the mid-20th century, mirroring shifts in cultural, technological, and social dynamics. The Billboard Hot 100 charts, established in 1958, have served as a barometer for musical popularity in the United States, chronicling the most popular genres, artists, and styles over the decades.

Leveraging extensive data from the Billboard charts and genre information from Spotify, this project aims to provide an examination of genre trends over time.

## 1.2 Data

The primary dataset for this project comprises records from the Billboard Hot 100 charts, spanning from 1958-2023. This dataset was sourced from a publicly available GitHub repository maintained by user HipsterVizNinja. It encapsulates the weekly rankings of songs in the United States, reflecting their commercial success and popularity. The raw dataset is structured as a CSV file titled “Hot\_100s.csv” located in the “raw-data” folder within the project repository.

To enrich this dataset with genre information, a cross-reference with the SpotifyAPI was performed. This process involved querying the SpotifyAPI for each unique artist to retreive their associated musical genres. The SpotifyAPI provides a comprehensive database of songs, artists, and their genres, making it an invaluable recourse for this analysis.

The data also underwent several cleaning and preprocessing steps to ensure its suitability for analysis. These steps included parsing and formatting dates, condensing the dataset for more efficent processing, and categorizing genres into broader classifications to simplify the analysis of genre trends.

## 1.3 Research Question

This project is guided by two primary questions:

1. **Genre Popularity Evolution:** How has the popularity of different music genres evolved over time according to the Billboard Hot 100 charts from 1958-2023?
2. **Audio Feature Evolution:** How have trends in audio features evolved over time across different music genres?

## 1.4 Purpose

The goal of this project is to uncover temporal trends in popular music and audio features in order to gain a better understanding of the shifting preferences of music listeners over time.

# 2. Methods and Results

## 2.1 Data Cleaning and Processing

The data cleaning and processing involved several key steps to prepare the dataset for analysis. The primary dataset utilized the Billboard Hot 100s charts records from 1958-2023, enriched with genre information via the SpotifyAPI. The process is as follows:

1. **Initial Preparation:** Removed unnecessary columns and renamed variables for clarify.
2. **Date Formatting:** Converted chart dates to a usable format and added “year”, “month”, and “week” columns for detailed analysis.
3. **Ranking and Filtering:** Established a ranking system based on chart position and duration, focusing on top songs and removing duplicates to streamline the dataset.
4. **Spotify API Integration:** Implemented the SpotifyAPI to fetch song and artist IDs, genre clasifications, and audio features.
5. **Genre Simplification:** Consolidated extensive genre labels into broader categories for analysis. Manual intervention was applied to assign genres to artists where Spotify data was missing.
6. **Data Merging and Finalization:** Merged enriched unique song data back into the main dataset, ensuring all entries contained compelte genre and audio feature information.
7. **Saving the Final Dataset:** The final processed dataset was saved to “processed-data” for analysis.

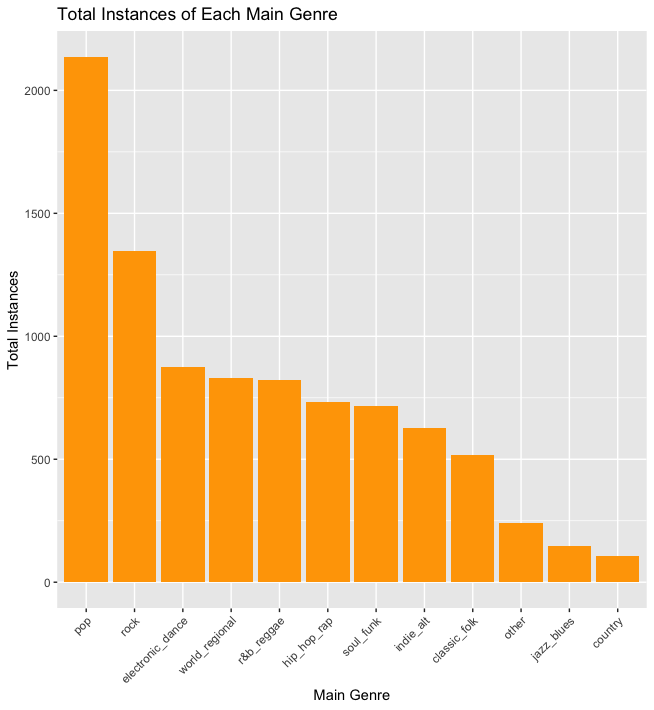
## 2.2 Exploratory Analysis

### 2.2.1 Summary Statistics

Summary statistics were generated for the dataset in order to gain an overarching understanding of its structure.

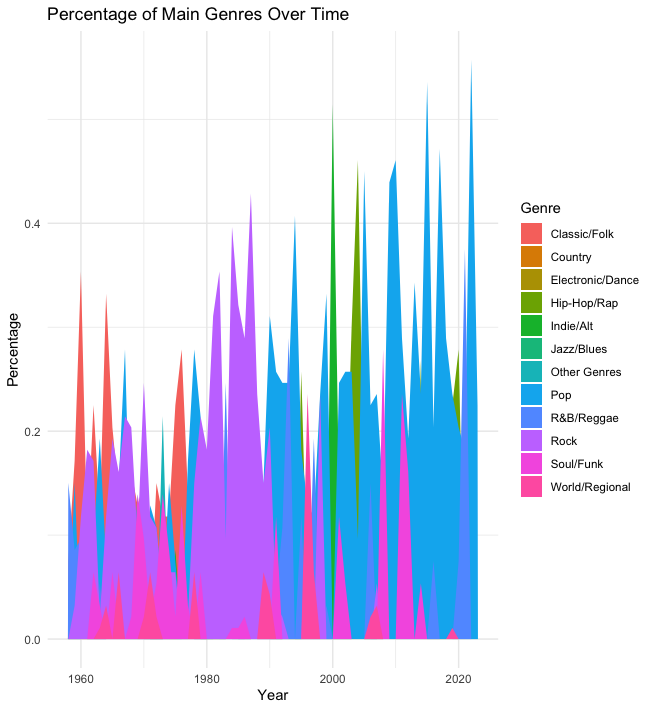
### 2.2.2 Genre Distribution

To begin to gain an idea of genre distribution, a bar graph was constructed to visualize the number of total instances for each main genre.



### 2.2.3 Genre Trends Over Time

In order to gain an idea of the genre trends over time, a rough visualization was constructed. The percentages of instances per year for each main genre were calculated, and then plotted into a chart.



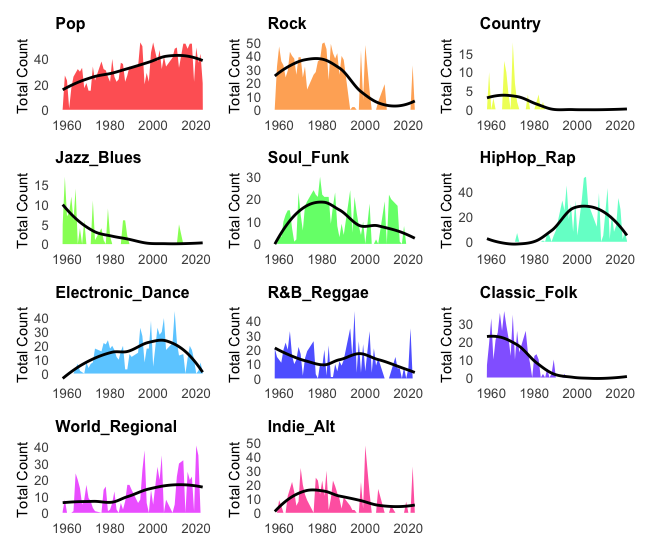
# 3. Analysis

## 3.1 Visual Analysis of Genre Popularity Over Time

This segment of analysis focused on the temporal trends in genre popularity. First, the dataset was loaded and prepared by converting main genre information into binary flags to facilitate analysis.

#### 3.1.0.1 Data Preparation and Visualization

After processing, the total number of songs per genre each year was calculated. The trends were visualized using area plots and smoothing lines to depict the changing popularity of genres over time. The use of the patchwork package allowed the plots to be stitched together in order to provide a comprehensive visual representation of genre dynamics.



## 3.2 Visual Analysis of Audio Features Over Time with Generalized Additive Models (GAMs)

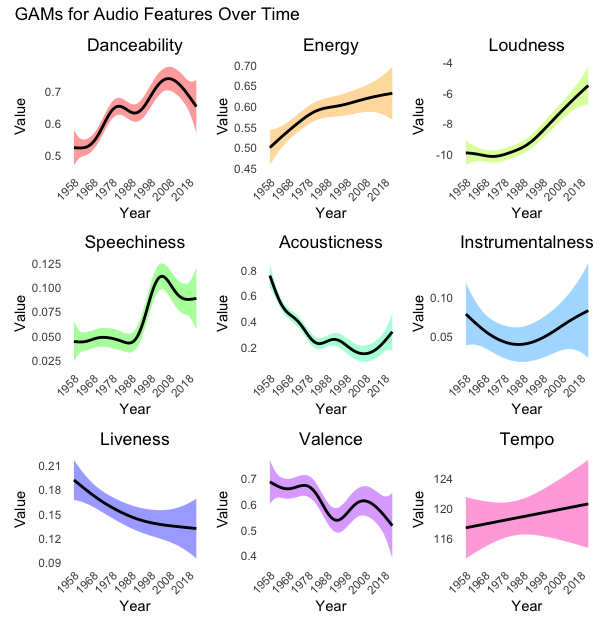
In an effort to understand the data further, Generalized Additive Models (GAMs) were utilized to analyze the relationship between audio features and time, adjusted for genre. This technique allowed for a more nuanced view of music evolution over time.

#### 3.2.0.1 Data Transformation and Modeling

The dataset was transformed to better suit a modeling approach, this was done by creating binary indicators for each main genre. For each audio feature, a GAM was fitted to capture non-linear trends and interactions over time, while considering the influence of genre.

#### 3.2.0.2 Visualization of Model Outputs

The outputs of the models were visualized using smooth trend lines in order to showcase the evolution of audio features over time. Each plot was color-coded to distinguish between features. Use of the patchwork package allowed for the plots to be combined into a unified grid which allowed for a more comprehensive view of audio features over time.



# 4. Conclusion

* Summarize the key findings of the project.
* Discuss the implications of these findings in the context of the research question.
* Suggest areas for future research or study.

# 5. References

* List all sources referenced throughout the project.