Decline of Rock Music in the 2000’s

(COMP3125 Individual Project)

Joshua Flores   
*School of Computing and Data Science*

*Abstract*—This research aims to study the fall in popularity of rock music at the turn of the century onwards. Using data from Spotify, Billboard, industry sales records, and the social media analytics three factors are examined, the rise in popularity of other genres, the transition from physical media to digital media, and the shift of bands to solo artists/singers. Results indicate a sharp decline in popularity of rock music following the rise of hip hop and pop music while also revealing how streaming platforms and social media have transformed how music is consumed in the modern era.

Keywords— music industry analysis, rock music, digital transformation, streaming platforms, genre evolution

# Introduction

The music industry has gone through many changes since the turn of the century, particularly in how it has been consumed, distributed and marketed. For the longest time music was typically an expensive endeavor with a vinyl record costing around 10 dollars in the 1980s [1]. This would come out to roughly 37 dollars in the present day. There is not much room for variety in music taste if nearly 40 dollars is dedicated to a singular collection of songs by 1 band. This is a stark difference to the current day where nearly every song that ever existed can be accessed for free with Ads through a streaming site or completely ad free with a subscription cost. This research focuses on rock music’s trajectory during the technological shift that has taken place in the last 24 years, examining multiple factors that have influenced its popularity and commercial success.

The study is particularly relevant to the current time as it coincides with the technological shifts in music consumption, from physical mediums to streaming platforms, and the rise in social media’s influence on music promotion and artist discoverability. A rise in other genres is observed as these shifts take place. This study aims to provide analysis on the combination of these impacts rather than their individual impacts.

# Datasets

## Source of dataset

1. Spotify Top Songs (2000-2023): Comprehensive data of popular songs on Spotify [2]

This was generated by the creator by taking the most popular songs on Spotify based on the year the song was released. Essentially the highest rated songs for each year were extracted and put into this data set, this explains the fact that Spotify did not release until 2006 this is simply measuring popular songs based on when they were released as older songs have a tendency to find popularity long after their heyday. [3]

1. Music Industry Sales Data (2000-2023): Format based sales figures from the recording industry [4]

This data set is pulled directly from tableau from the RIAA, Recording Industry Association of America, which a trade organization that represents the music recording industry in the United States. They have a crucial role in the legal and commercial landscaped of the music industry of the United States, so this data set is very credible.

1. Billboard Top 100 Data [5]

This data was pulled directly from the billboard website. Billboard is a leading American music industry magazine that tracks music charts, news, and trends in the music industry. This data is specifically the year end hot 100 for the years 2005, 2010, 2015, and 2020. The goal was to track which songs were by bands and which songs were by solo artists/ singers.

1. Social Media Usage Data: Platform Adoption rates from 2009-2023

This data tracks the popularity and adoption of social media sites throughout the years. The dataset itself is from Kaggle but the person that gathered the data used statcounter a website that allows users to track the market share of websites compared to their competitors. [7] The goal of this dataset was to help show the popularity increase in certain sites namely being youtube and Instagram.

## Character of the datasets

The Spotify dataset includes features such as genre classification, yearly rankings, popularity and various other columns to describe what type of song it is. As mentioned previously the data set is not the most popular songs on Spotify on each of the years listed rather it is the most popular songs on Spotify based on the year they were released.

The music sales data tracks revenue across different formats some being CDs, digital downloads, and streaming. This data set was very helpful to visualize just how drastic the shift from CD’s to digital medium was in the 2000s

The billboard data set categorizes artists as either bands or solo performers, tracking their chart positions over time. It is important to note that performances or songs where the singer clearly has a band that plays with them that this does not count as a band for these purposes. If the artist markets themselves as an individual they count as a singer/ solo artists regardless of if they have a live band they tour with.

The social media set was rather odd to visualize as it an abundant number of social media sites many of which were excluded from the visualization only the most important and relevant ones were kept for this research.

Each data sets, if possible, were filtered from the year 2000 onwards as that is the time period that is being focused on.

# Methodology

## Genre Analysis

For the genre analysis the Spotfiy data set was used and visualized the percentage of the top songs based on their genre. The visualization was a line chart showing the growth or decline of genres as time went on This would be the image for question 1 in the graphs directory. Specifically, the genres were narrowed down to “Rock, Hip Hop, Pop, and Electronic”. There was way too many genres if all were included. A mask was used to filter out the different genres as needed in the panda dataframe, all entries with the word rock were aggregating together. The line chart looks a little cluttered but it overall does the job.

## Sales Format Analysis

The analysis of music consumption formats utilized time series to track the transition from physical to digital formats. A stacked bar visualization was used to demonstrate the relative market share of different format of sales each year. The visualization makes the shift very clear. A key on the right hand side illustrates the color of the sale format for readability.

## Artist Type Analysis

To analyze the shift between bands and solo artists the number of top 100 songs that were made by singers vs by bands were plotted against each other to observe any pattern. Another visualization with the same data set was filtering out only the top 25 songs from each year in the data set 2005, 2010, 2015, 2020 and then creating a bar chart representing the number of songs that were by bands and by singers with the goal of seeing how the band number changed.

## Predictive Analysis

Two predictive models were created for the above data sets.

For the music sales data a random forest regressor model was used for projecting what the format distribution of sales would look like for the year 2023.

For the band vs singers data set a linear regression model was used to predict what rank a song by a band would be ranked in the year 2023 as songs by bands are few and far between in recent years.

# Results.

## Genre Evolution

Looking at the graph for Question 1 one can see that the red line for rock music is the highest out of all the genres right at the beginning on the year 2000. This number increases and peaks around the year 2006. However a few years after that one can see that pop begins to rise and then surpasses rock. In fact one may notice that pop actually explodes in popularity around 2010 while rock more or less flounders and stagnates. Rock never actually reaches the peak it sees in 2006 and eventually gets to a point where no rock song is actually a top song in the late 2019 to 2021. This visualization more or less shows a nose dive in the popularity of the genre while pop music sky rockets, meaning that factors such as digestibility and accessibility likely influence the genres fall and pop’s rise.

## Format Transition

Observing the visualization for Question 2 one can see that from 2000 to 2007 the amount of 8 track and CD purchases making up the sales were the majority of purchases by a very large margin. However, as time goes on the CD purchases get lower and lower each year until they are less than a quarter of all of the formats for that year. Starting from a place where CD’s and 8 tracks were 90% and above to less than a quarter is quite shocking not but surprising when one takes into account the rise in streaming platforms. In 2012 a small grey bar appears which represents those streaming platforms, streaming platforms are actually represented by several of the features on this visualization but this one is the most important. As time goes on that grey bar has increased until it is over 60% of the revenue generated for music. It is important to note that the overall revenue has technically gone down compared to the very start of the graph which can be attributed to the internet as a whole but also to the lack of want to own music physically.

*Predicted Distribution for 2023*

Using the data that was used to generate the previous graph a random forest regression model was used to predict what a bar on the graph may look like for the year 2023. This results in a graph similar to the 2019 bar with a much bigger grey bar which I am not inclined to disagree with because music consumption is only getting easier. However, an ever so slight increase in some of the physical music products may be seen and that is due to newer bands shifting towards selling them as collectors items as a way to make revenue but also be able to sell their fans something special.

## Band vs Solo Artist Trends

Observing the visualizations for the third question one can, maybe expectedly so, see that the number of songs in the top 100 by a band starts rather underwhelmingly and only continues to go down. On the other hand the number of songs in the top 100 for those years that include a singer start high and actually increase in number almost as a mirror. The nature of this graph makes sense as 2 directly opposite measures are being pitted against each other but it is still rather interesting to see. I expected the 2005 data to at least have half of the songs be by bands and it was almost there around 33 but it was already too close to the semi recent era of pop stars which exploded in popularity in the 2000s.

Another visualization was created using the same data however for this one the top 25 songs were filtered and the amount of songs by bands and singers were pitted against each other. The same trend can be seen singers increase in popularity and bands decrease. However, at the very beginning in 2005 the distribution is actually pretty close 9 songs out of the top 25 are made by bands.

*Predicting what rank a song by a band would have in 2023.*

A linear regression model was created on the bands vs singer data to try to predict what rank a song by a band would have in 2023 or if there would even be one at all. 80 percent of the data was taken as training and 20 percent was used for testing. The final results when predicting the rank for a song by any band (not a singer ) was 55.55 or 56. Meaning that if there was a song by a band in the top 100 for the year 2023 it would likely be around halfway up the ranking, nowhere near the top.

# Discussion

I would have liked to study more in depth the affect that these factors had on rock and all of its sub genres but since the scene is relatively niche it is rather difficult to find an abundance of data on it. If some of the data sets were crafted a little better I would have liked to use more complicated machine learning techniques such as Support Vector Machines or Neural Networks to draw relations in the data. In future work I would definitely be focused on finding more quality data and giving myself more time to really dive into the subject matter. I had many questions I was interested in tackling at the beginning, too many even, I think narrowing down on ones that are feasible and answerable will yield better results the next time I try this.

# Conclusion

In this part, you should summarize your project. What important results did you find for your topic and what’s the effect of this result on the real-world?

This research examined the decline of rock music in the 21st century, revealing 3 interconnected trends that have transformed the musical landscape a clear inverse relationship between rock and pop music popularity, with rock experiencing a dramatic decline after 2006 while pop music flourished. Then a shift from physical formats to streaming services changing how audiences consume music and a marked transition from bands to solo artists/singers in popular music charts. These results demonstrate how technological advancements and changing consumer preferences have reshapes the music industry. They initiated an environment that factors individual artists and streaming-optimized formats over traditional rock bands, marking an end to rock’s commercial dominance in popular culture.

##### References

1. "How Much Vinyl Records Cost in the 80s." *Vacation Vinyl*, Vacation Vinyl, <https://vacationvinyl.com/how-much-vinyl-records-cost-in-the-80s/>. Accessed Dec. 2024.
2. Van Eden, Conor. *Best Songs on Spotify for Every Year (2000-2023)*. Kaggle, 2023, <https://www.kaggle.com/datasets/conorvaneden/best-songs-on-spotify-for-every-year-2000-2023>.
3. "Tears for Fears' 'Everybody Wants to Rule the World' Tops Billboard Chart." *Billboard*, 6 Dec. 2024, <https://www.billboard.com/pro/tears-for-fears-everybody-wants-to-rule-the-world-tops-billboard-chart/>.
4. RIAA1295. *U.S. Recorded Music Revenues by Format*. Tableau Public, 2023, <https://public.tableau.com/app/profile/riaa1295/viz/U_S_RecordedMusicRevenuesbyFormat_0/RevenuesbyFormat>.
5. Billboard. *Billboard Charts*. Billboard, <https://www.billboard.com/charts/>. Accessed Dec. 2024.
6. Michau96. *Social Media Popularity (2009-2023)*. Kaggle, 2023, <https://www.kaggle.com/datasets/michau96/social-media-popularity-2009-2023>.
7. StatCounter. *StatCounter Global Stats*. StatCounter, <https://gs.statcounter.com/>. Accessed Dec. 2024.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**