Title:- Global Melodic Canvas- Insights into Top Songs

Student Name: Avinash Angilikam

Student ID: 23037971

Git Hub Link: https://github.com/AVINASH-ANGILIKAM/Global-Melodic-Canvas-Insights-into-Top-Songs

Abstract:

The report offers a comprehensive analysis of globally renowned songs over several decades. It covers various attributes such as artist details, song titles, release years, sales figures, streaming statistics, download metrics, radio airplay, and ratings. Through extensive analysis and visualization techniques, it reveals significant trends and patterns in popular music consumption. The findings provide valuable insights into the evolution of music preferences and the broader dynamics within the global music industry.

Introduction:

This report delves into the dynamic music industry landscape, examining top songs across several decades. Utilizing a comprehensive dataset including sales, streams, downloads, radio plays, and ratings, our objective is to uncover trends and patterns in popular music consumption over time. We aim to understand audience preferences and artist influence by investigating the impact of cultural, technological, and societal changes on music trends. Through data analytics and visualization techniques, we aim to unravel the intricate dynamics shaping the evolution of the music landscape, providing valuable insights into the constantly evolving trends and preferences in music.

Objective :

The "Global Melodic Canvas" report aims to analyze a music dataset to uncover key insights about the industry. Using visualizations and statistical analysis, we aim to understand trends in sales, streaming, artist prominence, and other important factors. Our goal is to provide practical insights for stakeholders in the music industry.

Descriptive Statistics Report:

Summary Statistics for Numerical Variables:

Artist: 4850 unique artists are represented in the dataset.

Title: There are also 4850 unique titles present.

Year: Data spans from 1901 to 2014.

Total Sales: Sales range from 4.149 to 36.503, with a mean of 8.154 and a median of 6.438. Standard deviation is approximately 4.556. The distribution is positively skewed (skewness = 2.017) and leptokurtic (kurtosis = 4.868).

Total Streams: Streams range from 0 to 25.545, with a mean of 4.629 and a median of 4.200. Standard deviation is approximately 3.274. The distribution is positively skewed (skewness = 1.134) and slightly leptokurtic (kurtosis = 2.281).

Total Downloads: Downloads range from 0 to 19.780, with a mean of 2.970 and a median of 2.200. Standard deviation is approximately 2.669. The distribution is positively skewed (skewness = 1.530) and leptokurtic (kurtosis = 2.852).

Total Radio Plays: Radio plays range from 0 to 24.393, with a mean of 3.985 and a median of 2.570. Standard deviation is approximately 4.354. The distribution is positively skewed (skewness = 1.572) and slightly leptokurtic (kurtosis = 2.518).

Total Rating: Ratings range from 0 to 4.460, with a mean of 0.482 and a median of 0.512. Standard deviation is approximately 0.625. The distribution is positively skewed (skewness = 1.716) and leptokurtic (kurtosis = 3.256).

Correlation Matrix: The correlation matrix reveals positive correlations among sales, streams, downloads, and radio plays. The strongest correlation exists between downloads and sales (0.662), followed by sales and radio plays (0.585).

Exploratory Data Analysis (EDA)

1.Bar Chart: Top 50 Artists by Sales

The bar chart visualizes the top 50 artists based on their sales. Each bar represents an artist, and the height of the bar corresponds to their sales figure. The bars are color-coded in light green with a black border for better visibility.
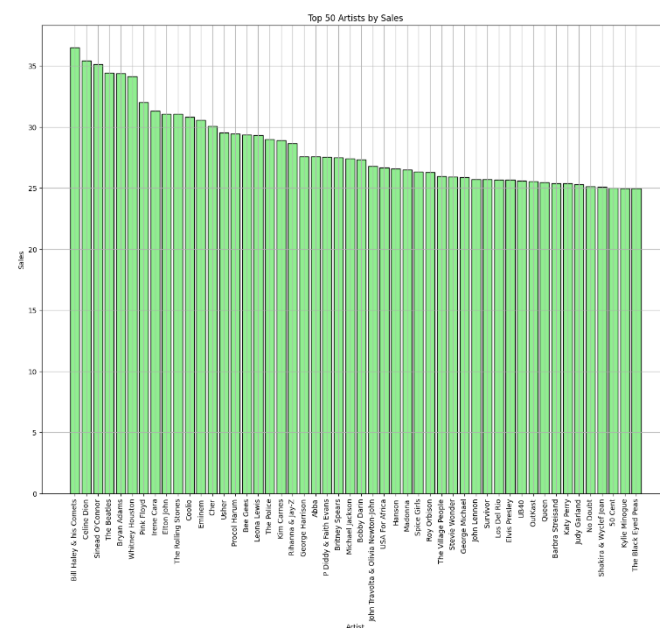

The x-axis represents the artists' names, and the y-axis represents their sales. To accommodate the artist

names without overlap, the x-axis labels are rotated by 90 degrees for better readability. This graph provides a quick overview of the top-performing artists in terms of sales, allowing for easy comparison and identification of the most successful artists within the dataset.

2.Line Graph: Sales, Streams, Downloads, and Radio Plays Up to 2014

The line graph below illustrates the trends of sales, streams, downloads, and radio plays from 1901 to 2014:

Sales, Streams, Downloads, and Radio Plays Up to 2014
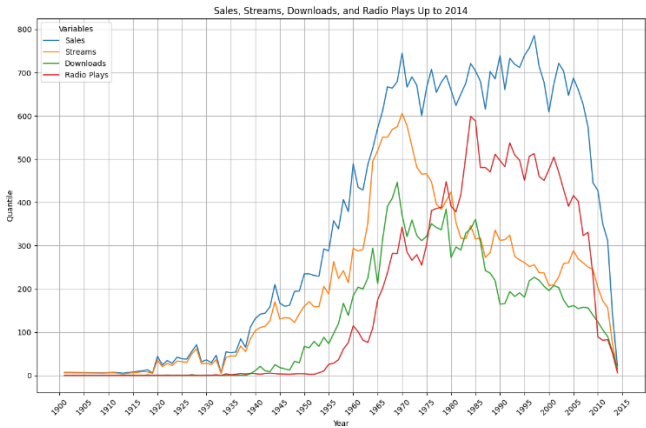
The x-axis represents the years from 1901 to 2014, while the y-axis represents the total quantile of each variable. Sales, streams, downloads, and radio plays are plotted over time, showing their respective trends and variations.

Key observations from the analysis reveal dynamic trends in music consumption patterns. Sales exhibit fluctuations, peaking from the 1960s to the late 1990s before gradually declining. Streaming patterns indicate a significant rise in popularity, particularly from the early 2010s, highlighting the increasing adoption of streaming services.

Similarly, downloads follow a similar trajectory as sales but experience a recent decline, potentially due to the shift towards streaming platforms. Despite digital advancements, radio plays remain relatively stable, serving as a significant channel for music discovery and promotion. These insights provide valuable understanding of evolving consumer behaviours and industry dynamics within the music landscape.

3.Correlation Heatmap

The correlation heatmap visualizes the pairwise correlations between numerical variables in the dataset. Each cell in the heatmap represents the correlation coefficient between two variables, with values ranging from -1 to 1.

Key findings from the analysis highlight significant correlations between various metrics within the music dataset.

Correlation Heatmap
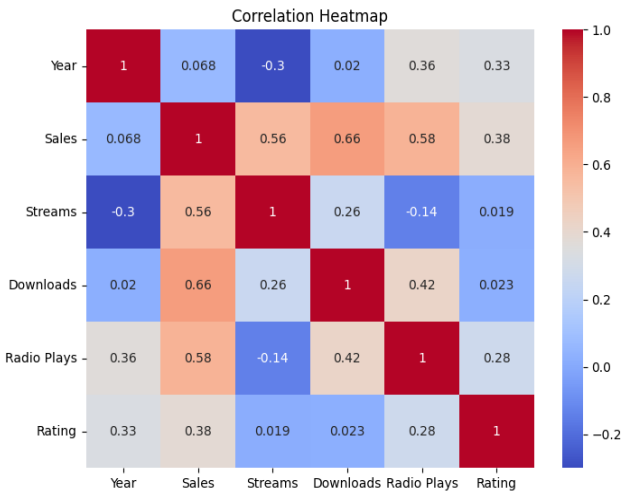
A strong positive correlation (0.662) is observed between sales and downloads, indicating that songs with higher download counts tend to have higher sales figures. Similarly, there is a strong positive correlation (0.556) between sales and streams, suggesting that songs with higher streaming numbers are likely to have higher sales. Additionally, radio plays exhibit a moderate positive correlation (0.584) with sales, indicating that songs receiving more airplay tend to have higher sales. Interestingly, a negative correlation (-0.300) is noted between the year of release and streams, implying that newer songs are streamed less compared to older ones. These findings offer valuable insights into the interplay between different metrics and their impact on the success of songs within the music industry landscape.

Conclusion:

The report offers a detailed overview of the dataset, including statistical summaries, variable correlations, and notable observations. It highlights potential trends and relationships that merit further investigation using advanced techniques.

The analysis of top songs globally offers key insights into the music industry dynamics. Market trends reflect shifts in consumer behaviours, notably influenced by digital platforms and characterized by a long-tail distribution. Rating disparities highlight differences in song quality and popularity, with successful tracks correlating to increased sales, downloads, and radio plays. These findings inform strategic decisions for stakeholders, guiding marketing, curation, and investment in navigating the music industry landscape.

Suggestions for Further Analysis:

1.Investigate the factors contributing to the observed skewness and kurtosis in the distributions.

2.Explore potential outliers and their impact on the data analysis.

3. Conduct additional analyses to understand the underlying factors influencing sales, streams, downloads, radio plays, and ratings.