

# **Analyzing Spotify's Dataset Using Power BI**

**LUCENO, CRISTOPHER JOHN**  
DATA ARCHITECTURE, DATA 039-008  
Continuing Education and Professional Studies  
Southern Alberta Institute of Technology

## **TABLE OF CONTENTS**

- I.** Introduction
- II.** Dataset Overview
- III.** Objectives
- IV.** Key Questions
- V.** Metrics & KPIs
- VI.** Entity Relationship Diagram (ERD)
- VII.** Data Analysis & Insights
- VIII.** Visualizations & Findings
- IX.** Conclusion

## I. INTRODUCTION

In the contemporary digital landscape, streaming platforms have significantly altered how audiences engage with music. Spotify, one of the world's leading audio streaming services, offers a vast catalog of over 100 million tracks, shaping the way listeners interact with their favorite artists and genres. Given this dynamic transformation, analyzing streaming data provides invaluable insights into audience behavior, music trends, and industry patterns. This study aims to leverage Power BI for an in-depth analysis of Spotify's most streamed songs from 2020 to 2023, uncovering trends, artist popularity, and the characteristics that define successful tracks. Through data visualization and advanced analytics, I will explore how music consumption patterns evolve over time and what factors contribute to a song's widespread appeal.

## II. DATASET OVERVIEW

The dataset utilized for this analysis is sourced from Kaggle, <https://www.kaggle.com/datasets/abdulszz/spotify-most-streamed-songs>, and comprises a comprehensive collection of data on Spotify's most streamed songs. It includes essential attributes such as song names, artists, total stream counts, release years, and various audio features like danceability, tempo, and energy levels. By methodically analyzing this dataset, I seek to extract key insights regarding the factors that influence song popularity. This dataset serves as the foundation for our study, allowing us to examine the impact of various characteristics on streaming numbers and uncover potential correlations between audio features and listener preferences.

## III. OBJECTIVES

The primary objective of this research is to conduct an extensive analysis of Spotify's most streamed songs and determine the fundamental elements that contribute to their success. This study aims to:

- Identify the core factors that influence a song's streaming performance and longevity.
- Examine streaming trends over time, assessing the evolution of musical preferences among listeners.
- Evaluate how playlists, charts, and distinct audio features contribute to a song's popularity and virality. By utilizing Power BI, I aim to develop an interactive and intuitive representation of these insights, providing a data-driven approach to understanding modern music consumption patterns.

## IV. KEY QUESTIONS

To guide our exploration, I have identified several key questions that will be addressed through my analysis:

1. What are the top 10 most streamed songs from 2020 to 2024, and what factors contributed to their success?
2. Who are the most streamed artists, and how does their streaming performance compare across different years?
3. How do streaming numbers vary across different genres, and which genres have seen the highest growth?
4. What overarching trends can be observed in music popularity over time, and how do external factors influence these trends?
5. How do various song characteristics, such as danceability, tempo, energy, and loudness, correlate with the total number of streams? These questions form the backbone of my analysis, ensuring a structured and focused approach to extracting meaningful insights from the dataset.

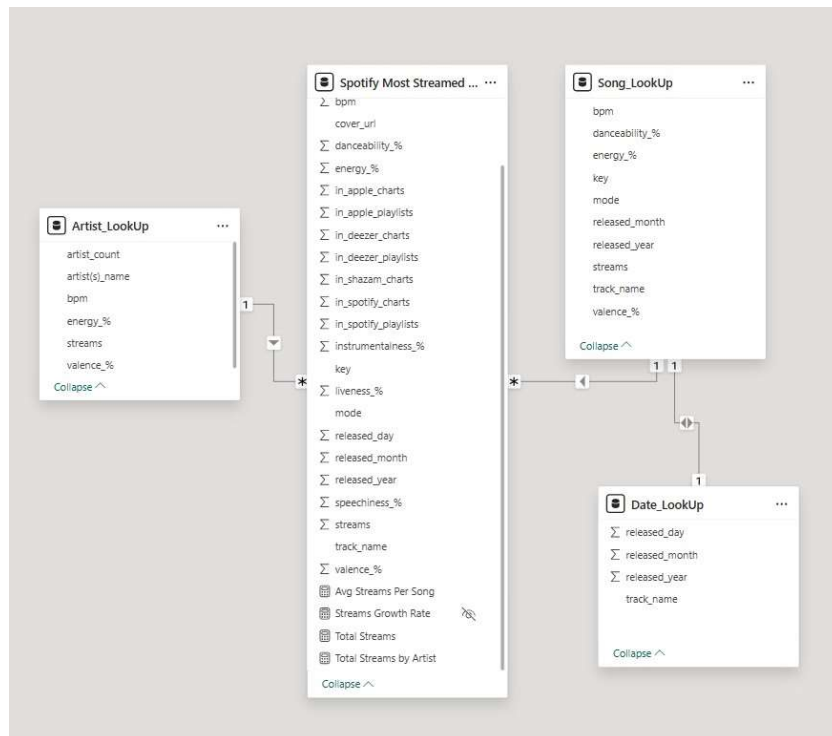
## V. METRICS & KPIs

To effectively measure and assess the success of my analysis, I have identified a set of key performance indicators (KPIs) that provide a structured framework for evaluation. These include:

- **Total Streams (2020-2023):** The cumulative number of streams recorded for each song over the analysis period, providing insights into overall popularity.
- **Average Streams per Song (2020-2023):** The mean number of streams across all songs, highlighting the general engagement levels across different tracks.
- **Most Streamed Artists (2020-2023):** Identification of top-performing artists based on cumulative streams, offering a comparative analysis of artist success.
- **Correlation Between Audio Features and Streams (2020-2023):** An analytical breakdown of how various song attributes, such as tempo, energy, and danceability, influence a song's streaming numbers. By evaluating these KPIs, we can establish patterns that define successful songs and provide actionable insights for artists and industry stakeholders.

## VI. ENTITY RELATIONSHIP DIAGRAM (ERD)

To ensure a structured approach to data modeling, I developed an Entity Relationship Diagram (ERD) that visually represents the relationships between various components of our dataset. The ERD helps establish connections between different attributes, such as songs, artists, genres, and stream counts, ensuring an organized and efficient data structure. This model allows for a more streamlined analysis, enabling me to extract insights with greater accuracy and coherence.



## VII. DATA ANALYSIS & INSIGHTS

Our analysis, conducted using Power BI, involved a thorough examination of streaming trends by categorizing data based on release year, genre, and artist. The key findings from our analysis reveal:

- **Genre Dominance:** Pop and hip-hop genres consistently dominate the most-streamed songs list, underscoring their strong listener base.
- **Influence of Audio Features:** Songs that exhibit higher danceability and energy levels tend to garner more streams, suggesting that lively and engaging tracks resonate more with audiences.
- **Artist Performance Trends:** A select group of artists consistently appears in the top-streamed lists, indicating their strong industry influence and fan loyalty.

- **Fluctuating Streaming Patterns:** Streaming numbers are influenced by various external factors, including the release of new albums, social media trends, and playlist placements. These insights provide a deeper understanding of music consumption trends, helping industry professionals make data-driven decisions.

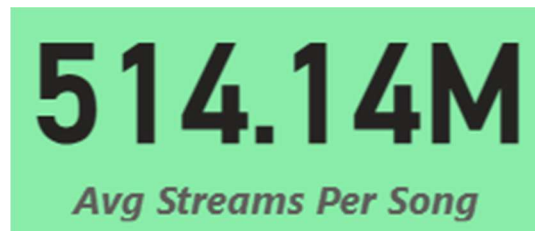
## VIII. VISUALIZATION & FINDINGS

To effectively communicate our findings, we utilized Power BI to create various data visualizations, including:

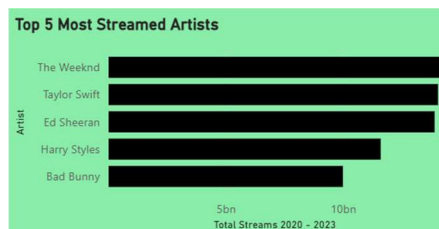
- **Sum of Total Streams from 2020-2023:** A cumulative total of all streams recorded during this period, providing a high-level view of overall music consumption.



- **Average Streams per Song from 2020-2023:** An analytical view of the average stream count per song, offering insights into general listening behavior.



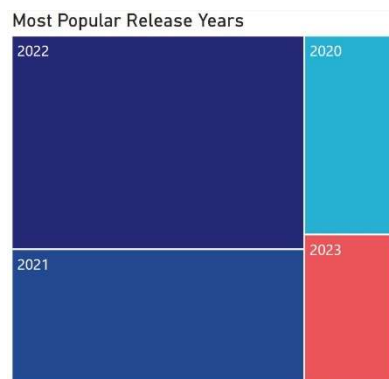
- **Top 5 Most Streamed Artists:** A comparative bar chart showcasing the top five artists based on total streams, reflecting their influence and popularity.



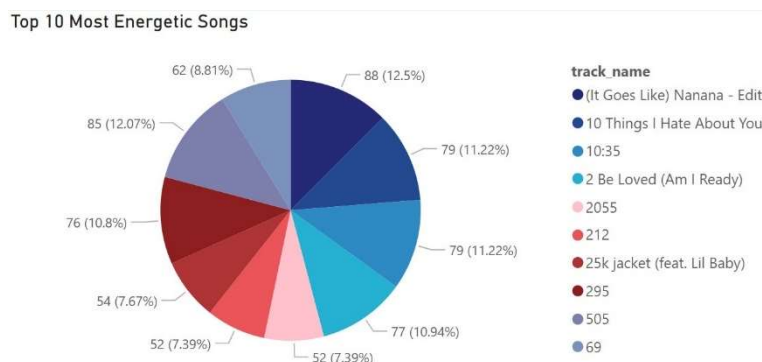
- **Top 5 Most Streamed Songs:** A ranked list of the five highest-streamed songs, providing insight into listener preferences.

Top 5 Most Streamed Songs						
Artist	Blinding Lights	Dance Monkey	Shape of You	Someone You Loved	Sunflower	Total
The Weeknd	3703895074					3703895074
Ed Sheeran			3562543890			3562543890
Lewis Capaldi				2887241814		2887241814
Tones and I		2864791672				2864791672
Post Malone, Swae Lee					2808096550	2808096550
<b>Total</b>	<b>3703895074</b>	<b>2864791672</b>	<b>3562543890</b>	<b>2887241814</b>	<b>2808096550</b>	<b>15826569000</b>

- **Most Popular Release Years:** A breakdown of the most common release years among the top streamed songs, illustrating how song age impacts streaming performance.



- **Top 10 Most Energetic Songs:** A visualization ranking the ten songs with the highest energy scores, showing the correlation between energetic music and streaming numbers. These findings offer valuable perspectives into the dynamics of music streaming, helping us understand how various factors influence song popularity.



## **IX. CONCLUSION**

Through this comprehensive analysis, we successfully transformed raw Spotify data into meaningful insights using Power BI. We identified critical streaming trends, assessed artist dominance, and explored the impact of audio features on song popularity. This study highlights the significance of data visualization in understanding music consumption and industry dynamics, demonstrating how analytical tools can provide valuable business intelligence for artists, record labels, and music executives.