**Exploring Music Trends: A Deep Dive into Spotify Data**

**Author:** Lakshmi Sreeja Gotika

**Date:** 05-06-2024

# **Table of Contents**

1. Introduction
2. Dataset Description
3. Methodology
4. Data Analysis and Visualization
5. Conclusion
6. Appendix
7. **Introduction**

**Objective**

The primary objective of this project is to analyze music trends using Spotify data. By leveraging a comprehensive dataset that includes various attributes such as track names, artist names, release dates, and performance metrics like streams and chart positions, this project aims to uncover key insights into music consumption patterns. The analysis focuses on identifying key performance indicators (KPIs) such as total streams, average danceability percentage, and the total number of tracks. These KPIs will help in understanding the popularity and characteristics of different tracks and artists.

**Scope**

This project covers the following areas:

* Data Collection and Preparation:
* Key Performance Indicator (KPI) Analysis
* Data Visualization
* Insight Generation

#### **Significance**

Understanding music trends is crucial for various stakeholders in the music industry, including artists, record labels, streaming platforms, and marketers. This project provides valuable insights into:

* Popularity of Tracks and Artists
* Music Consumption Patterns
* Characteristics of Popular Music
* Impact of Collaborations

This project not only provides a snapshot of current music trends but also offers a foundation for predicting future trends in the music industry. The insights gained from this analysis can inform decision-making for artists and industry professionals, helping them to strategize their releases and marketing efforts more effectively.

1. **Dataset Description**

**Source:** Kaggle

* **track\_name: Name of the track.**
* **artist(s)\_name: Names of the artists.**
* **artist\_count: Number of artists involved.**
* **released\_year: Year the track was released.**
* **released\_month: Month the track was released.**
* **released\_day: Day the track was released.**
* **in\_spotify\_playlists: Number of Spotify playlists featuring the track.**
* **in\_spotify\_charts: Number of Spotify charts featuring the track.**
* **streams: Number of streams on Spotify.**
* **in\_apple\_playlists: Number of Apple playlists featuring the track.**
* **in\_apple\_charts: Number of Apple charts featuring the track.**
* **in\_deezer\_playlists: Number of Deezer playlists featuring the track.**
* **in\_deezer\_charts: Number of Deezer charts featuring the track.**
* **in\_shazam\_charts: Number of Shazam charts featuring the track.**
* **bpm: Beats per minute.**
* **key: Musical key.**
* **mode: Musical mode.**
* **danceability\_%: Danceability percentage.**
* **valence\_%: Valence percentage.**
* **energy\_%: Energy percentage.**
* **acousticness\_%: Acousticness percentage.**
* **instrumentalness\_%: Instrumentalism percentage.**
* **liveness\_%: Liveness percentage.**
* **speechiness\_%: Speechiness percentage.**

### **Methodology**

### **Data Collection**

### The dataset for this project was sourced from Spotify, one of the largest and most popular music streaming platforms globally. The dataset comprises various attributes related to tracks, artists, and performance metrics, which are essential for conducting a thorough analysis of music trends. The specific steps involved in data collection were as follows:

### **Data Cleaning**

### To ensure the accuracy and reliability of the analysis, the dataset underwent a rigorous cleaning process:

### **Handling Missing Values:**

### Identified and addressed missing values by either imputing them with appropriate values or excluding incomplete records if necessary.

### **Data Type Correction:**

### Ensured that each attribute had the correct data type (e.g., numerical, categorical) to facilitate accurate analysis.

### **Duplicate Removal:**

### Removed duplicate entries to prevent skewed results, ensuring each track and artist appeared only once in the dataset.

### **Normalization:**

### Standardized values, particularly for textual data such as artist names and track titles, to maintain consistency across the dataset.

### 

### **Data Analysis**

### The cleaned dataset was then subjected to various analytical techniques to extract meaningful insights:

### **Descriptive Statistics:**

### Calculated basic statistics (mean, median, mode, standard deviation) for numerical attributes to understand the distribution and central tendencies.

### **Correlation Analysis:**

### Examined relationships between different attributes (e.g., danceability and streams) to identify potential patterns and influences.

### **Grouping and Aggregation:**

### Aggregated data based on different dimensions (e.g., by artist, by year) to facilitate comparative analysis.

### **Visualization**

### Visual representations of data are crucial for uncovering trends and patterns. The following visualizations were created using Excel and Qlik Sense:

### **KPI Charts:**

### **Total Streams:**

### Displayed as a numerical value highlighting the overall popularity of the tracks.

### **Average Danceability Percentage:**

### Showed the average danceability score, indicating how danceable the tracks are on average.

### 

### **Track Count:**

### Illustrated the total number of tracks in the dataset, providing a sense of the dataset's scope.

### **Bar Chart:**

### Visualized streams for each track, allowing for easy comparison of track popularity.

### **Pie Chart:**

### Represented the distribution of total streams among different artists, highlighting the most influential artists.

### **Line Chart:**

### Depicted the number of streams over time, showing trends in music popularity across different years.

### **Storytelling**

### To present the analysis effectively, a structured storytelling approach was used, divided into five slides:

### 

### 

### 

### 

### 

### 

#### 

1. **Data Analysis and Visualization**

#### **KPI Analysis**

The analysis began by identifying and calculating key performance indicators (KPIs) from the dataset:

* Total Streams: Summed up the streams for all tracks, highlighting overall popularity.
* Average Danceability Percentage: Calculated the average danceability score across all tracks, indicating how suitable they are for dancing.
* Track Count: Counted the total number of unique tracks, providing a sense of the dataset's breadth.

These KPIs were visualized in a consolidated KPI chart, providing a quick overview of the dataset's performance metrics.

#### **Bar Chart Analysis**

A bar chart was created to visualize the number of streams for each track. This chart allowed for easy comparison of track popularity and highlighted the top-performing tracks. Notable findings include:

* Tracks with the highest streams, such as "Seven (feat. Latto)" and "LALA".
* Comparison of single-artist versus collaborative tracks and their streaming performance.

#### **Pie Chart Analysis**

The distribution of total streams among different artists was represented using a pie chart. This visualization provided insights into the dominance of certain artists and their influence on streaming trends:

* High stream counts for artists like Jung Kook, Myke Towers, and Olivia Rodrigo.
* The impact of popular artists on the overall streaming numbers.

#### 

#### **Line Chart Analysis**

A line chart was used to depict the number of streams over time, focusing on the release year of tracks. This visualization illustrated trends in music popularity across different years:

* Identification of peak years for streaming, with 2023 showing significant activity.
* Analysis of trends showing increasing or decreasing popularity over time.

#### **Snapshots**

Each chart was supplemented with snapshots and annotations to highlight significant points, making the visualizations more informative and easier to interpret.

### **Conclusion**

This project aimed to analyze music trends using a dataset from Spotify, focusing on key performance indicators (KPIs) such as total streams, average danceability percentage, and track count. Through detailed data analysis and visualization, several significant insights were uncovered:

1. **Popularity Trends:**

* Tracks with higher streams often featured multiple artists, indicating the strong appeal of collaborations in the music industry.
* Recent years, particularly 2023, showed an increase in the number of streams, reflecting a growing trend in music consumption via streaming platforms.

1. **Artist Influence:**

* Dominant artists like Taylor Swift, Bad Bunny, and Jung Kook played a significant role in the overall streaming landscape, with their tracks consistently appearing at the top of the charts.
* The influence of established artists with large fan bases was evident, as they contributed a substantial share of the total streams.

1. **Track Characteristics:**

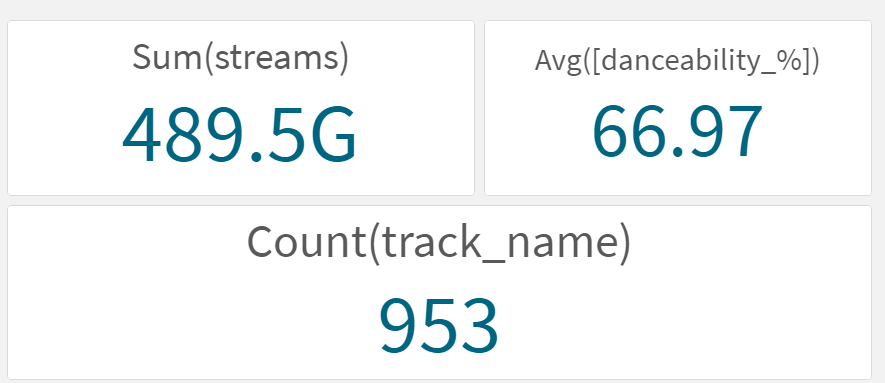
* Popular tracks generally had higher danceability and energy scores, suggesting that listeners prefer engaging and upbeat music.
* The analysis of musical attributes like BPM, key, and mode provided additional insights into the characteristics of tracks that resonate with audiences.

These findings offer valuable information for various stakeholders in the music industry, including artists, producers, and marketers. By understanding the factors that drive track popularity and the influence of artist collaborations, industry professionals can better strategize their releases and marketing efforts. This project also highlights the importance of continuous data analysis to keep up with evolving music consumption trends.

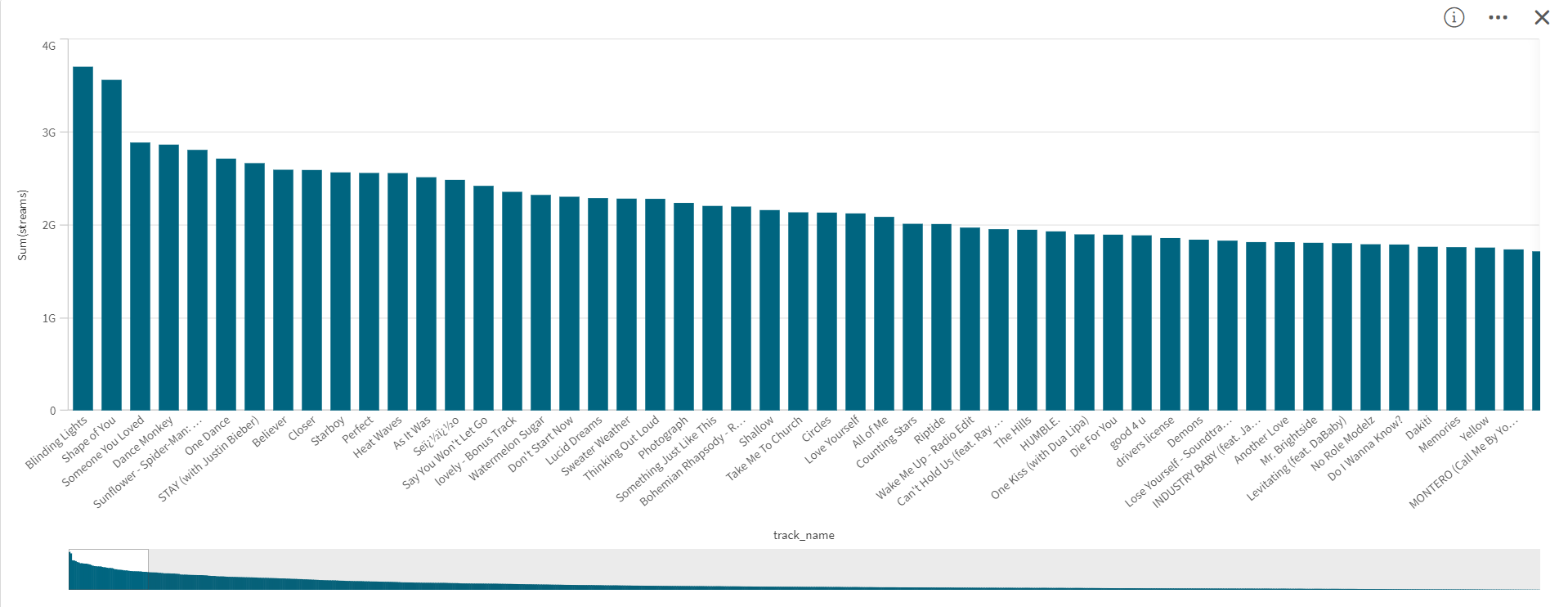
**6. Appendix**

#### **Additional Charts:**

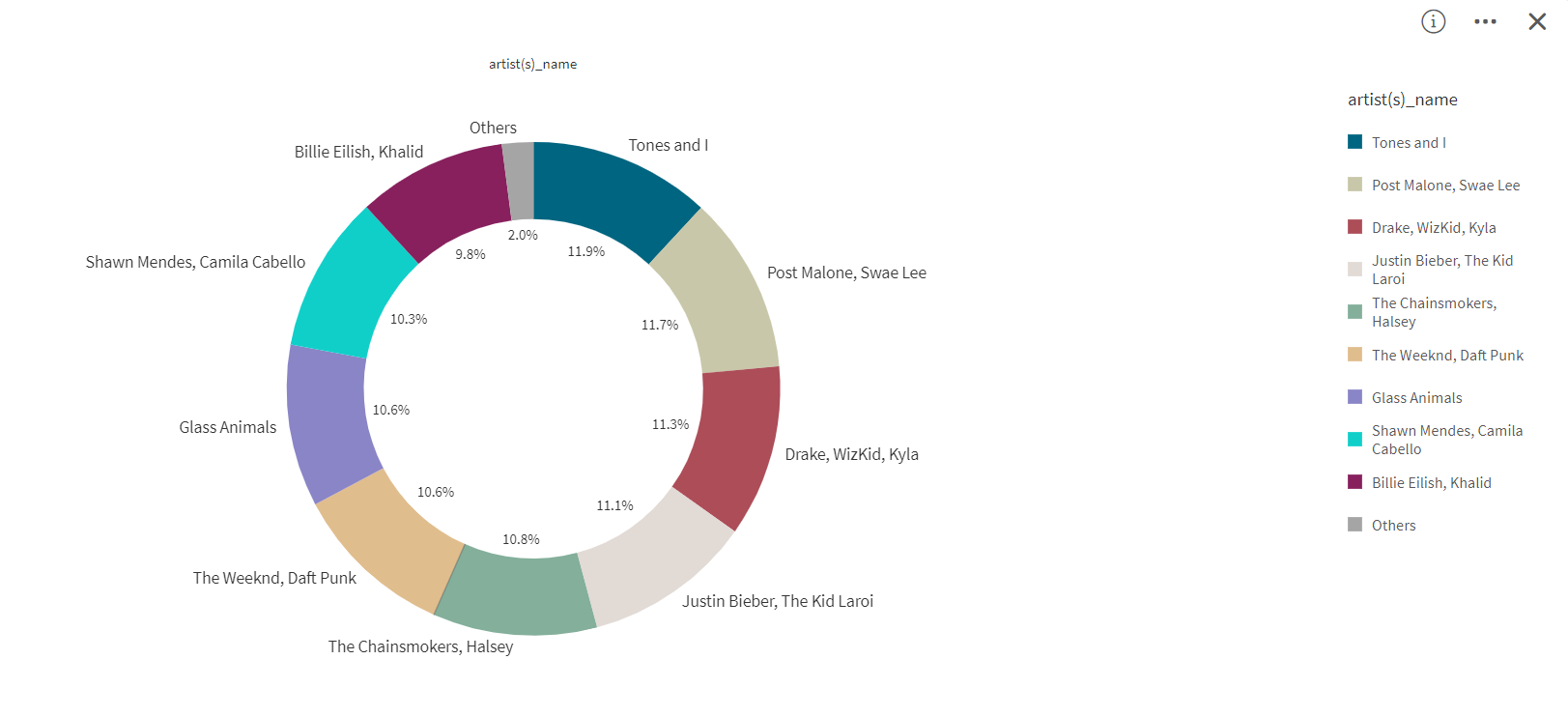
* **KPI Charts**

****

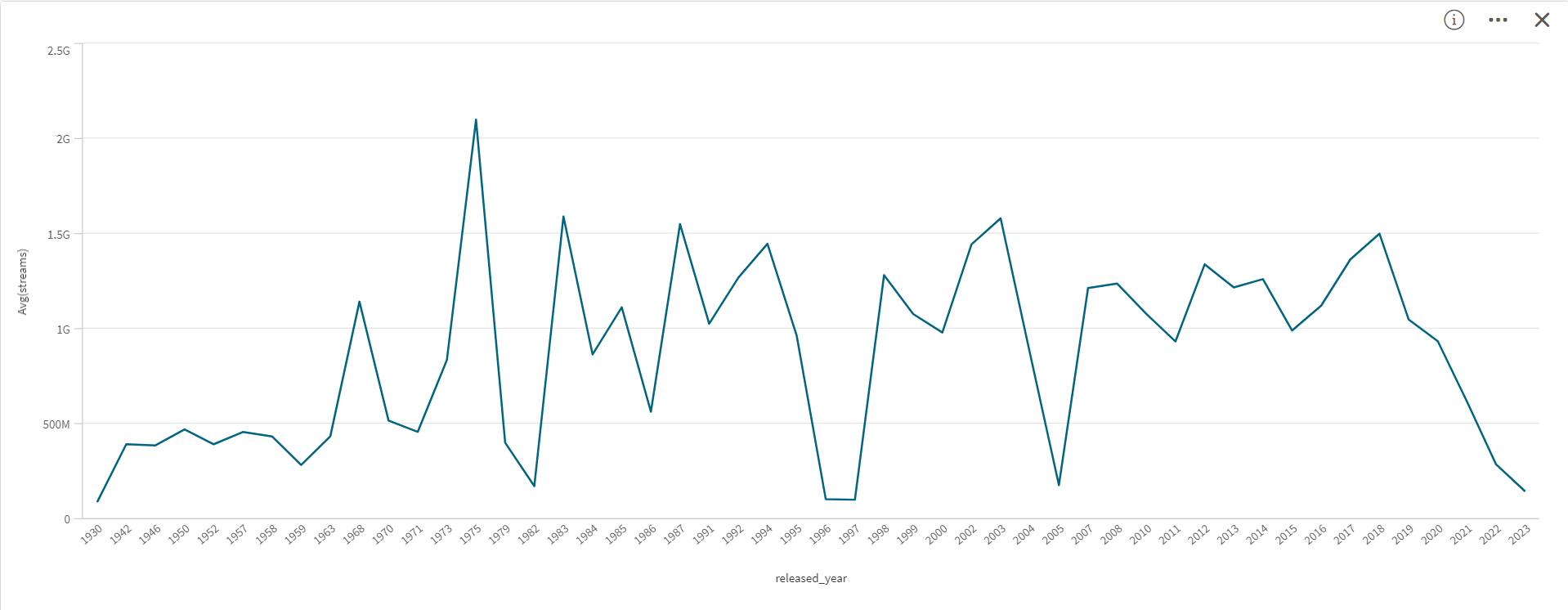
* **Bar Chart**

****

* **Pie chart**

****

* **Line Chart**

****