****

**Spotify Data Analysis**

**Sanjana Dalvi**

**Table of Contents**

* Introduction
* Data Source
* Data Import
* Dashboard Explanation
* KPI Analysis
* Charts Analysis
* Conclusion

**Introduction**

The **Spotify Data Analysis Dashboard** provides an interactive visualization of Spotify streaming data. The goal of this project is to analyze the performance of various artists and tracks, identify trends over time, and examine musical attributes like danceability, energy, and speechiness. The dashboard helps users quickly derive insights from complex datasets for better understanding and decision-making in the music industry.

**Data Source**

**Data Source:**

The data for this project was sourced from **Kaggle**, a widely-used platform for datasets and data science projects. The dataset contained detailed information on Spotify tracks, including:

* **Artist Name**
* **Track Name**
* **Streams Count**
* **Release Year**
* **Musical Attributes** such as tempo, danceability, energy, and speechiness.

**Data Import**

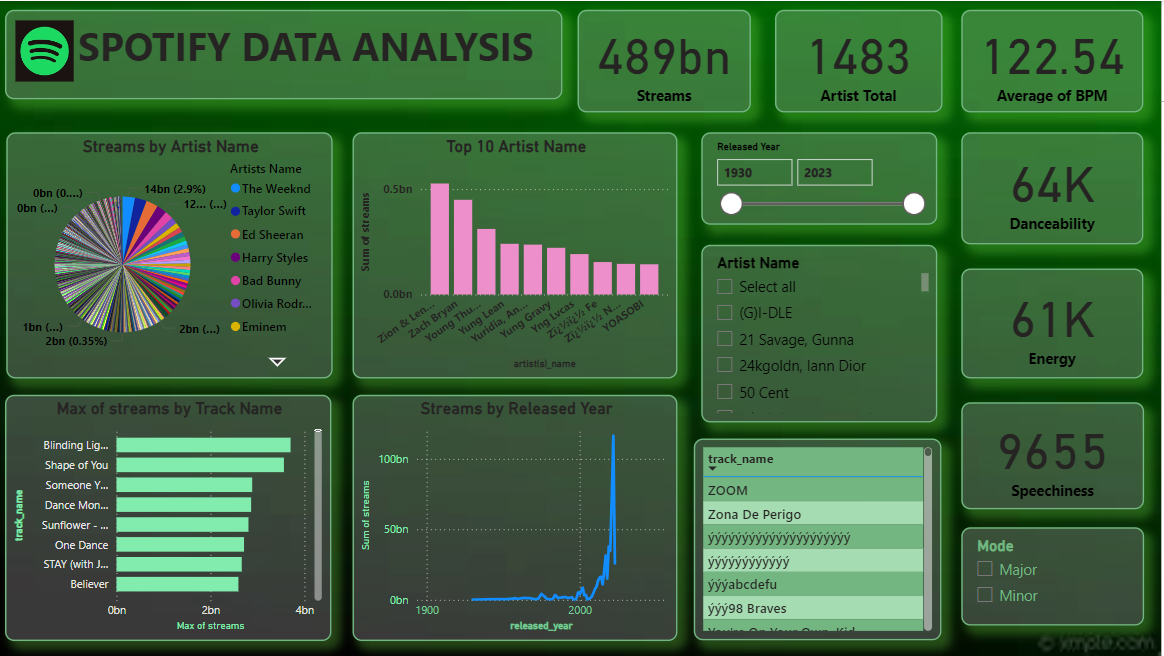
**Loading Data into Power BI:**

* + The dataset, which was in CSV format, was imported into Power BI by navigating to the **Home** tab and selecting **Get Data** > **Text/CSV**. Once selected, the file was loaded into Power BI, and the initial preview of the dataset was displayed.
  + After clicking **Load**, the dataset became available for transformation and visualization in Power BI.

**Data Cleaning in Power BI:** Using Power BI’s **Power Query Editor**, several cleaning steps were performed to prepare the dataset for analysis:

* + **Removing Duplicates:** In the **Power Query Editor**, any duplicate rows were identified and removed using the **Remove Duplicates** feature to ensure the dataset only contained unique records.
  + **Handling Missing Data:** Missing or null values in critical fields such as **Streams**, **Danceability**, and **Energy** were addressed. The missing values were either removed or replaced with default values using the **Replace Values** function:
    - For numerical columns, missing values were imputed with the mean or median.
    - In categorical columns (like Artist Name), missing values were replaced with **"Unknown"**.

**Dashboard Explanation**

****

Visualization: Multiple visuals were created, including pie charts, bar charts, line graphs, and card visuals for total streams, artists, and other key metrics. Each charts explain in next page.

**KPI Analysis**

The top panel of the dashboard highlights key aggregate metrics:

* **Total Streams:** 489 billion streams, representing the dataset's scale.
* **Total Artists:** 1483 unique artists analyzed.
* **Average BPM (Beats Per Minute):** 122.54, indicating the tempo of songs on average.
* **Danceability:** 64,000 (out of a normalized scale), reflecting how suitable tracks are for dancing.
* **Energy:** 61,000, showing the intensity or activity level of the tracks.
* **Speechiness:** 9655, measuring the presence of spoken words in the music.

**Charts Analysis**

**Pie Chart: Streams by Artist Name**

* A pie chart visualizing the distribution of streams across different artists.
* The legend lists artists like **The Weeknd**, **Taylor Swift**, **Ed Sheeran**, and **Harry Styles**, showing their percentage of total streams.
* The chart demonstrates that **The Weeknd** has the largest share of streams at 2.9%, with other artists contributing varying amounts.

**Bar Chart: Top 10 Artist Name**

* A bar chart representing the top 10 artists based on total stream count.
* Artists like **Zion & Lennox**, **Bryson Tiller**, and others have high numbers of streams, with the exact count displayed in billions.

**Bar Chart: Max of Streams by Track Name**

* A bar chart that shows the top tracks based on their maximum stream count.
* Popular tracks such as **Blinding Lights**, **Shape of You**, and **Someone You Loved** have the highest streams, with **Blinding Lights** leading at nearly 4 billion streams.

**Line Graph: Streams by Released Year**

* A line graph that shows the trend of total streams over the years.
* The graph shows a sharp rise in streams after the year 2000, reflecting the increasing popularity of digital music streaming.

**Release Year Slider**

* A slider that allows the user to filter the data based on the release year of tracks, from 1930 to 2023. This interactive feature lets users explore data for specific time periods.

**Artist Name (Checklist)**

* A filter that allows the user to select one or more specific artists for analysis. The user can select from a list of popular artists like **(G)I-DLE**, **21 Savage**, **Gunna**, and more.

**Mode**

* A filter showing the musical mode of the tracks, either **Major** or **Minor**. This indicates the general mood of the music, with major often being more upbeat and minor more somber.

**Conclusion**

This Power BI dashboard allows users to explore and analyze Spotify streaming data efficiently. The use of Power BI’s data transformation tools enabled the creation of a clean and insightful data model, with interactive visuals providing valuable insights into artist performance, track popularity, and musical trends over time.

**Thank You**