Data Visualization and Analysis of Most Streamed Spotify Songs (2024)

1. Project Overview

This project aims to analyze the most streamed songs on Spotify in 2024 to uncover insights about artist popularity, streaming trends, and explicit content distribution. The dataset provides valuable information about track performance across different platforms, helping to identify key patterns in music streaming trends.

2. Dataset Description

The dataset contains information about the most streamed songs on Spotify in 2024. It includes the following key columns:

- Track – Name of the song.
- Artist – Name of the performing artist.
- Album Name – The album in which the song was released.
- Release Date – The date the song was released.
- Spotify Streams – Total number of streams on Spotify.
- Spotify Playlist Count – Number of playlists the track is included in.
- Spotify Playlist Reach – The reach of the track on Spotify playlists.
- YouTube Views & Likes – Engagement metrics from YouTube.
- TikTok Posts & Views – The song's presence and engagement on TikTok.
- Explicit Track – Indicates whether the track contains explicit content (Yes/No).

3. Data Cleaning Process

Before performing any analysis, we applied the following data cleaning steps:

a. Handling Missing Values

- ‘Soundcloud Streams’ column had 72.46% missing values – this was removed from the analysis.
- ‘TIDAL Popularity’ column had 100% missing values – this was dropped.

- Other missing values were handled through imputation or removal based on data relevance.

b. Removing Duplicates

- Checked for duplicate entries using data.duplicated().sum().
- Removed duplicate records to maintain data integrity.

c. Converting Data Types

- The 'Release Date' column was converted to a datetime format for better time-based analysis.
- Categorical columns were properly encoded where needed.

4. Exploratory Data Analysis (EDA)

We performed data visualization and analysis to derive key insights:

a. Top Artists by Spotify Streams

- Created a bar chart showcasing the most streamed artists.
- Artists with the highest streams were identified and ranked.

b. Number of Tracks Released Over the Years

- A bar graph was used to visualize trends in song releases.
- This helped in understanding the volume of music production per year.

c. Explicit vs. Non-Explicit Tracks

- A bar chart comparing the count of explicit vs. non-explicit songs.
- This provided insights into the nature of trending songs.

d. **Correlation Analysis**

- Used a heatmap to identify relationships between various streaming metrics.
- Found significant correlations between Spotify streams and YouTube engagement.

5. Conclusion

- The analysis revealed that a few top artists dominate streaming numbers.

- TikTok popularity significantly influences Spotify streams, showing a strong link between social media and streaming success.
- The majority of trending songs are non-explicit, but explicit tracks still hold a large share.

This project successfully analyzed streaming trends and provided actionable insights into the music industry's dynamics in 2024. Further analysis can explore sentiment analysis on lyrics and deeper machine learning models for stream predictions.