

## Introduction

This report presents an analysis of a Spotify dataset, focusing on track popularity, duration, explicit content, and other characteristics. The objective is to uncover insights into the trends and patterns within the music available on Spotify. The analysis was performed using Python, with libraries such as pandas, numpy, matplotlib, and seaborn for data manipulation and visualization.

## Methodology

**Data Loading:** The dataset, named `tracks.csv`, was loaded into a pandas DataFrame for analysis. This dataset includes various attributes of tracks on Spotify, such as name, popularity, duration, and artists.

## Data Exploration and Cleaning

Initial exploration was conducted to understand the dataset's structure.

Missing values were identified and analyzed.

Data types were examined and adjusted as necessary, with special attention given to the `release\_date` column, which was converted to datetime format.

## Data Analysis Techniques Applied:

**Descriptive Statistics:** Summary statistics were generated to provide an overview of the data's central tendencies and variability.

**Sorting and Sampling:** The dataset was sorted based on track popularity, and a sample was taken to facilitate detailed analysis.

**Correlation Analysis:** A correlation heatmap was created to explore relationships between numerical variables.

**Trend Analysis:** The distribution of tracks over years and the change in track duration over time were analyzed and visualized.

## Analysis Findings

**1. Track Popularity:** The analysis identified tracks with the highest popularity, providing insights into current music trends.

**2. Duration Trends:** There has been a noticeable variation in track duration over the years, with recent trends and historical comparisons highlighted.

**3. Correlation Insights:** A correlation analysis between variables such as loudness and energy, and popularity and acousticness, revealed significant relationships that can influence a track's success and listener preferences.

**4. Genre Popularity:** The analysis extended to genres, showcasing the duration of songs across different genres and identifying the most popular genres based on track popularity.

## Visualization Highlights

The report includes several visualizations:

A heatmap for correlation analysis.

Scatter plots and regression plots to illustrate relationships between variables.

Histograms and line plots to show the distribution and trends of track releases over the years.  
Bar plots to compare the duration and popularity of songs across different genres.

### **Conclusion**

The Spotify dataset analysis revealed insightful trends and patterns in music preferences and track characteristics. Popularity is closely related to factors like acousticism and energy, and there is a significant variation in track durations over the years. Additionally, genre analysis highlighted the varying popularity and music styles that resonate with listeners.