# **Report on Spotify Popularity Analysis and Predictive Modeling**

## **1. Introduction**

In this analysis, we explore the relationship between spotify’s popularity and various audio features such as danceability, energy, loudness, tempo, and genre. The goal is to uncover patterns and trends that can help predict the popularity of a song and understand which characteristics contribute to its success. The project is divided into four main objectives, each aiming to provide insights into different aspects of music popularity.

This report addresses the following objectives:

* **Objective 1:** Analyzing the Relationship Between Popularity and Audio Features
* **Objective 2:** Genre Analysis and Popularity Trends Over Time
* **Objective 3:** Identifying Characteristics of 'Hit' Songs
* **Objective 4:** Predicting Song Popularity Using Audio Features and Genre

Each section of the report provides a detailed explanation of the methods used, the findings, and the visualizations created to support the analysis.

## **Objective 1: Analyzing the Relationship Between Popularity and Audio Features**

### **Purpose**

The purpose of this objective is to investigate how different audio features—specifically danceability, energy, loudness, and tempo—affect the popularity of songs. By identifying the correlations between these features and song popularity, we aim to uncover which characteristics are most strongly associated with a song's success.

### **Methodology**

* **Correlation Matrix (Heatmap):** A correlation matrix was created to visualize the relationships between the popularity and other audio features. This heat map allows for a quick overview of the strength and direction of correlations.
* **Scatter Plots:** Scatter plots were generated to examine individual relationships between popularity and each audio feature.

### **Findings**

* The **correlation matrix** showed that danceability and energy have moderate positive correlations with popularity, suggesting that songs with higher groove and intensity tend to be more popular.
* The scatter plots confirmed that tracks with higher energy and danceability were more likely to receive higher popularity scores. This indicates that energetic and rhythm-driven tracks have a higher chance of becoming hits.

## **Objective 2: Genre Analysis and Popularity Trends Over Time**

### **Purpose**

This objective explores how the popularity of songs has evolved over time and how it varies across different genres. By analyzing trends in music taste and the rise and fall of genres, we aim to understand the shifts in consumer preferences and industry trends.

### **Methodology**

* **Line Chart:** A line chart was created to show the overall trend in popularity across the years, helping identify changes in music taste over time.
* **Stacked Bar Chart:** A stacked bar chart was used to analyze how the popularity of different genres has evolved over the years.

### **Findings**

* The **line chart** revealed an increasing trend in popularity over the years, indicating a general rise in music consumption or changing music preferences.
* The **stacked bar chart** showed that genres like pop and hip-hop have maintained high levels of popularity, while genres such as rock have seen a decline over time.

## **Objective 3: Identifying Characteristics of 'Hit' Songs**

### **Purpose**

In this objective, we identify the key characteristics that define 'hit' songs, focusing on their audio features (danceability, energy, loudness, tempo) and genre. The goal is to understand the traits that successful songs share and determine if there are any common patterns that could predict a song’s success.

### **Methodology**

* **Box Plots:** Box plots were used to examine the distribution of audio features for songs with low vs. high popularity scores, helping identify the characteristics that are more common in hit songs.
* **Bar Chart:** A bar chart was created to compare the average values of different audio features for hit songs across various genres.

### **Findings**

* **Box plot analysis** showed that hit songs generally have higher energy, danceability, and tempo. These features are distributed in the higher ranges for popular songs, suggesting that energetic, groovy tracks with faster tempos are more likely to be hits.
* The **bar chart** confirmed that genres like pop and hip-hop have higher average values for features like danceability and energy, making them more likely to produce popular tracks.

## **Objective 4: Predicting Song Popularity Using Audio Features and Genre**

### **Purpose**

The goal of this objective is to build a predictive model that can forecast song popularity based on its audio features and genre. By using machine learning techniques such as regression analysis or classification models, we aim to understand which features are most important for predicting popularity and to develop a model that can predict whether a song will become a hit.

### **Methodology**

* **Data Preprocessing:** The dataset was cleaned and processed to handle missing values and encode categorical variables like genre.
* **Modeling:** A **linear regression model** was used to predict continuous popularity scores, while a **classification model** was tested for binary prediction (hit vs. non-hit).
* **Feature Importance:** The models were analyzed to determine which features (e.g., energy, danceability, genre) contributed the most to predicting song popularity.

### **Findings**

* The regression model revealed that energy, danceability, and genre were the most significant predictors of popularity.
* The classification model was successful in predicting whether a song would be a hit or not, with features like energy and danceability playing key roles in the predictions.

## **Conclusion**

This analysis has provided valuable insights into the factors that influence song popularity. By examining the relationships between audio features, genre, and popularity, we have identified the key characteristics that define successful songs. The predictive models developed can assist in forecasting which songs are likely to become hits based on their features.

The findings highlight the importance of characteristics like **energy**, **danceability**, and **genre** in determining a song's popularity. These insights can be used by record labels, artists, and marketers to better understand musical trends and the factors that contribute to a song's success.