

# Music Data Analysis Assignment

## Overview

[This dataset](#) contains user-generated content (UGC) music uploads from 01/2023-01/2025 across multiple countries, limited to songs with total >30-second-plays (play30s) between 3,000 and 1,000,000. It includes engagement metrics such as downloads, favorites, reposts, and playlist ads. Your goal is to analyze this data to surface meaningful insights about song and artist performance.

*Note: The 'month' column represents the calendar month of activity (e.g., '4/1/2024' means April 2024, not a single day of play).*

## Objective

Use the dataset to identify:

1. **Early Indicators of Success** – What engagement signals (or combinations) are the strongest predictors of a song gaining traction?
2. **Trends & Patterns** – Identify any notable trends across countries, time periods, or user interactions.
3. **Hidden Gems** – Surface songs or artists that appear to be on the rise but may not be obvious based on total plays alone.
4. **Areas for Improvement** – Highlight any data gaps or limitations that might impact decision-making.

## Deliverables

- **Summary Report (1 page max)** detailing key insights and interpretations, with a short list of specific songs surfaced during your analysis.
- **SQL Queries** used to extract insights (clearly commented).
- **Data Visualization (Optional)** to support your findings (Excel, Python, or any tool of choice).

## Guidelines

- **Be creative** – there's no single "right" answer.
- **Justify your conclusions** – what makes a song promising beyond just play count?
- **Think practically** – if you were advising a music platform to better understand its content, or advising a label on who to sign, what would you recommend based on your findings?

## Submission Details

- **Format:** PDF report + SQL queries (.sql or .txt).
- **Deadline:** 02/28/2025.