Project Title: Spotify & YouTube Music Analytics Dashboard for Streaming Insights

Client Name: Harmony Media Solutions Pvt. Ltd.

Project Overview:
Harmony Media Solutions, a global leader in music streaming analytics, wants to enhance its data-driven decision-making capabilities by building an interactive Power BI dashboard. The client aims to analyze music streaming trends across Spotify and YouTube to gain insights into track performance, user engagement, and content effectiveness. The dashboard will provide real-time insights into popular songs, artist performance, and audience engagement metrics across both platforms.

Business Problem Statement:
The client, a music streaming analytics company, aims to understand how songs perform across Spotify and YouTube based on various attributes like danceability, energy, popularity, and engagement metrics (views, likes, comments). The insights will help music labels, artists, and marketing teams make data-driven decisions on promotions, collaborations, and content strategies.

Business Objective:
Develop an interactive Power BI dashboard to analyze song performance across Spotify and YouTube, helping stakeholders identify trends, engagement patterns, and key success factors for music streaming.
The primary objective of this project is to develop a **data visualization and analytics solution** that will enable Harmony Media Solutions to:

1. **Understand song performance** across Spotify and YouTube.
2. **Identify key trends** in user engagement based on metrics such as streams, views, likes, and comments.
3. **Compare artist and album performance** across different platforms.
4. **Analyze correlations** between song attributes (e.g., danceability, energy, valence) and popularity.
5. **Identify content strategies** that maximize reach and engagement for music labels and independent artists.

Dataset Overview:
Dataset of songs of various artist in the world and for each song is present:

- Several statistics of the music version on spotify, including the number of streams;
- Number of views of the official music video of the song on youtube.

The dataset consists of Spotify and YouTube statistics for various songs, including:

- Spotify Metrics: Streams, danceability, energy, acousticness, loudness, valence, tempo, and more.
- YouTube Metrics: Views, likes, comments, official video status, and channel information.
- General Information: Song name, artist, album, and release type

Fields:

It includes 26 variables for each of the songs collected from spotify. These variables are briefly described next:

- Track: name of the song, as visible on the Spotify platform.
- Artist: name of the artist.
- Url_spotify: the Url of the artist.
- Album: the album in wich the song is contained on Spotify.
- Album_type: indicates if the song is relesead on Spotify as a single or contained in an album.
- Uri: a spotify link used to find the song through the API.
- Danceability: describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.
- Energy: is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.
- Key: the key the track is in. Integers map to pitches using standard Pitch Class notation. E.g. 0 = C, 1 = C♯/D♭, 2 = D, and so on. If no key was detected, the value is -1.
- Loudness: the overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typically range between -60 and 0 db.
- Speechiness: detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks.
- Acousticness: a confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.
- Instrumentalness: predicts whether a track contains no vocals. "Ooh" and "aah" sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly "vocal". The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0.
- Liveness: detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live.
- Valence: a measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).
- Tempo: the overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.

- Duration_ms: the duration of the track in milliseconds.
- Stream: number of streams of the song on Spotify.
- Url_youtube: url of the video linked to the song on Youtube, if it have any.
- Title: title of the videoclip on youtube.
- Channel: name of the channel that have published the video.
- Views: number of views.
- Likes: number of likes.
- Comments: number of comments.
- Description: description of the video on Youtube.
- Licensed: Indicates whether the video represents licensed content, which means that the content was uploaded to a channel linked to a YouTube content partner and then claimed by that partner.
- official_video: boolean value that indicates if the video found is the official video of the song.

Note:
These datas are heavily dependent on the time they were collected, which is in this case the 7th of February, 2023.

Scope of Work:
The project will include the following key components:

1. Data Integration & Processing
- Extract and clean the dataset containing Spotify and YouTube music performance metrics.
- Handle missing or inconsistent data to ensure accuracy in reporting.
- Transform data into a structured format for Power BI visualization.

2. Developing Key Performance Indicators (KPIs):
- Top-performing artists based on streams and views.
- Correlation between Spotify popularity and YouTube engagement.
- Influence of song attributes (danceability, energy, loudness) on success.
- Most engaging YouTube videos (likes-to-views ratio, comments analysis).

3. Dashboard Development
- **Overall Music Performance Overview**
  - Total streams, views, likes, and comments by platform.
  - Top-performing artists and albums.
- **Track Analytics**
  - Key performance indicators (KPIs) for individual tracks.
  - Correlation between Spotify attributes (e.g., energy, valence, danceability) and streaming performance.
- **Artist Performance Insights**
  - Comparative analysis of artists based on engagement metrics.
  - Breakdown of official vs. non-official music videos.
- **Trend Analysis & Forecasting**

- Identification of rising trends in the music industry.
- Historical data visualization for strategic insights.

- **User Engagement Analysis**
  - Sentiment analysis based on comments and likes.
  - Analysis of how different content formats impact engagement.

4. Basic to Intermediate-Level DAX Calculations:

- Total streams/views across all songs.
- Like-to-view and comment-to-view ratios.
- Average energy, danceability, and valence for top songs.
- Year-wise trend analysis using time intelligence functions.
-

5. Insights & Recommendations

- Provide actionable insights for music labels and content creators.
- Suggest optimal strategies for content distribution and audience targeting.

Technical Requirements:

- **Power BI Desktop** for dashboard creation.
- **Excel** for data preprocessing (if required).
- **Local Storage** for data management (since cloud-based Power BI services are not available)

Expected Deliverables:

1. **Interactive Power BI Dashboard** with real-time insights on music streaming trends that allows stakeholders to analyze music trends across Spotify and YouTube, Visual representation of top-performing artists, songs, and engagement patterns.
2. **Data Cleaning & Preprocessing Documentation** outlining steps taken for data preparation.
3. **Analytical Report** summarizing key findings, correlations, and business recommendations, Insights on song attributes affecting engagement and popularity.
4. **Executive Summary** to provide high-level insights for stakeholders.
5. **Actionable recommendations** for music labels, marketers, and artists.

Project Timeline:

Day	Task
Day 1	Dataset exploration, understanding business goals, defining KPIs.
Day 2	Data cleaning, handling missing values, and transformations.
Day 3	Creating visualizations and setting up dashboard structure.
Day 4	Implementing DAX calculations and refining visuals.
Day 5	Adding interactivity (filters, slicers, drill-through).
Day 6	Dashboard finalization and optimization.
Day 7	Project presentation and client review.

Success Criteria:

- The dashboard should be **user-friendly** and **interactive**, providing actionable insights at a glance.
- Business users should be able to **filter and analyze data dynamically** based on artist, song, platform, and engagement metrics.
- The solution should help **identify streaming patterns and content strategies** to improve audience reach.

Constraints & Considerations:

- The project will be executed only in Power BI Desktop (no scheduled refresh via Power BI Service).
- The dataset reflects data collected on February 7, 2023, and is not live-updated.
- Advanced DAX is not required, only basic to intermediate calculations will be used.

Final Submission Requirements:

- Power BI .pbix file with complete dashboard and DAX calculations.
- Summary report with insights and key findings (optional).
- Project presentation explaining methodology and conclusions.

Conclusion:

This project will empower Harmony Media Solutions to make data-driven decisions in the evolving music industry landscape. By leveraging Power BI, the company can gain a competitive edge in understanding audience preferences and optimizing content strategies for better engagement and revenue growth.