Spotify Music Data Analysis Using SQL

Problem Statement

With the ever-growing volume of music released on streaming platforms like Spotify, understanding what drives a song's popularity has become crucial for artists, producers and marketers. This project aims to explore and analyze Spotify music dataset to uncover trends related to track popularity, artist performance and user engagement.

The core objective is to extract actionable insights by performing exploratory data analysis (EDA) and solving business-focused queries using SQL, leveraging aggregate functions, common table expressions (CTEs) and window functions.

Key Insights and Solutions

Track Popularity & Performance

- Identified tracks with 1B+ streams to highlight high user engagement.
- Retrieved top 5 high-energy tracks to reveal energetic content trends.
- Analyzed tracks with energy-to-liveness ratio > 1.2 for engagement cues.

Artist & Album Analysis

- Listed albums with respective artists to understand contribution spread.
- Counted total tracks by each artist, helping identify the most and least prolific contributors.
- Calculated average danceability per album, giving a metric for musical vibe consistency.

User Engagement

- Summarized views and likes for official videos, revealing user interaction levels.
- Found cumulative likes using window functions, allowing us to track growth over time.
- Identified top 3 most-viewed tracks per artist, useful for targeted promotions.

Business-Focused Insights

- Detected tracks with above-average liveness, indicating potential for live performances.
- Calculated total views per album to assess overall reach.
- Measured total comments on licensed tracks, reflecting listener feedback levels.