Spotify Music Trends & Popularity Analysis

Business Intelligence Insights Report

1. Introduction and Project Context

The music streaming industry is one of the most data-driven digital sectors, where competition is strongly influenced by the ability to understand listener preferences and predict content performance. Spotify, as a leading global music streaming platform, relies extensively on data analytics to optimize recommendations, curate playlists, and support both artists and business partners.

This Business Intelligence (BI) project aims to analyze Spotify music data to identify the key factors that drive song popularity. While Spotify collects vast amounts of information on tracks, genres, and artists, determining what truly contributes to a song's success remains complex. Popularity is influenced not only by genre and artist reputation, but also by intrinsic audio characteristics such as energy, danceability, and emotional tone (valence).

The objective of this analysis is to transform raw music data into actionable insights using a full BI pipeline: data preparation, dimensional modeling, dashboard development, and insight generation. The final outcome is an analytical foundation that helps explain music popularity trends and supports strategic decision-making related to content curation, playlist design, and artist promotion.

2. Data and BI Solution Overview

2.1 Data Description

The dataset used in this project contains over 12,000 music tracks released between 2015 and 2020. Each track is described using a combination of categorical and numerical attributes, including:

- Audio features (energy, danceability, valence, acousticness, instrumentalness)
- Track metadata (artist, genre, duration, explicit content)
- Time attributes (release year and date)
- Popularity score (ranging from 0 to 100)

This dataset provides a rich foundation for analyzing how musical characteristics and trends relate to popularity outcomes.

2.2 Data Modeling

To support efficient analysis and scalability, a **star schema** dimensional model was implemented. The central fact table, **FactTracks**, stores quantitative measures such as popularity and audio feature values. It is connected to four dimension tables:

- **DimGenre**: genre classification
- **DimArtist**: artist information
- **DimTrack**: track-level metadata
- **DimDate**: time hierarchy (year, quarter, month)

This modeling approach improves query performance, enables flexible slicing and filtering, and aligns with BI best practices.

2.3 KPIs and Measures

Several DAX measures were defined to capture performance and trends, including:

- Average Popularity
- Popularity Year-over-Year (YoY) and Month-over-Month (MoM)
- Average Energy, Danceability, and Valence
- Explicit Content Percentage
- Average Track Duration
- Total Tracks and Number of Artists

These measures enable both high-level performance monitoring and detailed exploratory analysis.

2.4 Dashboard Structure

The Power BI solution consists of three main dashboard layers:

1. **Executive Summary Dashboard**
   Provides an overview of total tracks, average track duration, number of artists, genre distribution, and popularity trends over time.
2. **Deep Dive Dashboard**
   Focuses on relationships between audio features and genre, and comparative analysis.
3. **Drill-Through Artist Dashboard**
   Enables detailed exploration of individual artist performance, track-level popularity, and historical trends.

Together, these dashboards support descriptive and diagnostic analytics, forming the basis for the insights discussed in the following section.

3. Key Findings and Business Interpretation

3.1 Popularity Has Increased Over Time

The analysis shows a clear upward trend in average song popularity from 2015 to 2020. This suggests that tracks released in more recent years tend to achieve higher popularity scores compared to earlier releases.

Why this matters:
This trend may reflect improvements in Spotify's recommendation algorithms, changes in listener behavior, or evolving music production standards. It also indicates that newer content is more competitive and that continuous content refreshment is essential to maintain engagement.

3.2 Track Duration Is Relatively Standardized

The average track duration across the dataset is approximately 3.46 minutes, indicating a strong standardization of song length in modern music production.

Why this matters:
Since most tracks follow a similar duration format, track length is unlikely to be a major driver of popularity. This shifts the focus toward audio features and genre characteristics as key differentiating factors.

3.3 Popularity Is Concentrated in a Small Number of Genres

Pop-related genres (Pop, Pop/Indie, Pop Rock, and Pop Ballad) dominate the catalog in terms of track count. They also show higher average popularity compared to many niche genres.

Why this matters:
While Spotify offers a diverse catalog, user attention and streaming volume are concentrated in a limited number of genres. This concentration highlights the importance of balancing mainstream appeal with catalog diversity.

3.4 Explicit Content Represents a Significant but Minority Share

Non-explicit tracks account for approximately 62% of the catalog, while explicit tracks represent about 38%.

Why this matters:
Explicit content is common but not dominant. Spotify must balance artistic freedom with accessibility, age restrictions, and regional content regulations.

3.5 Danceability Shows a Strong Positive Relationship with Popularity

Tracks with higher danceability scores tend to achieve higher popularity. The scatter analysis reveals a consistent upward pattern, indicating that rhythm and beat suitability play an important role in listener engagement.

Why this matters:
Danceable tracks are more likely to be included in playlists for workouts, parties, and social settings, which increases their exposure and replay rate.

3.6 Energy Is a Defining Feature of Popular Songs

Genres and tracks with higher energy levels tend to achieve higher popularity compared to low-energy tracks.

Why this matters:
Energy reflects intensity and excitement, which aligns with modern listening habits and supports higher engagement across playlists and recommendations.

3.7 Valence Has a Moderate but Positive Impact on Popularity

Valence, which represents musical positivity, shows a positive but weaker relationship with popularity compared to danceability and energy. Tracks with moderately positive emotional tones tend to perform better than very low-valence tracks.

Why this matters:
Listeners appear to prefer emotionally balanced or uplifting music, but extreme happiness is not necessarily a requirement for popularity. Emotional diversity remains important.

3.8 A Small Number of Artists Drive Consistent High Popularity

The artist drill-through analysis shows that a limited group of artists consistently achieves high average popularity across multiple tracks and years.

Why this matters:
Artist brand strength plays a major role in popularity. Successful artists contribute disproportionately to platform engagement and streaming volume.

3.9 Audio Profiles Differ Significantly Across Genres

Genre-level analysis reveals strong differences in average valence, danceability, and energy across musical genres.

Why this matters:
These differences indicate that recommendation and playlist strategies should be genre-aware rather than relying on uniform audio thresholds.

3.10 Certain Artists Show Long-Term Popularity Consistency

The drill-through analysis of BTS shows consistently high popularity across multiple years and tracks, with an average popularity above 70.

Why this matters:
Consistent artist performance indicates strong fan loyalty and brand value, making such artists key drivers of sustained platform engagement.

4. Business Recommendations

4.1 Use Clear Audio Rules When Building Playlists

Spotify should use simple audio rules (for example, minimum danceability and energy levels) when selecting tracks for playlists such as workout, party, or social playlists.

Why this helps:
This ensures playlists are consistently built around the audio features that are shown to increase popularity and listener engagement.

4.2 Adapt Recommendations to Each Genre's Characteristics

Spotify should adjust its recommendation logic based on the typical audio features of each genre instead of applying the same rules to all genres.

Why this helps:
Different genres have different musical structures, so genre-aware recommendations feel more relevant and personalized to listeners.

4.3 Identify High-Potential New Artists Using Audio Features

Spotify should use audio features such as energy, danceability, and valence to identify new or less-known artists whose tracks have strong popularity potential.

Why this helps:
This allows Spotify to promote promising artists earlier, instead of relying only on past popularity or artist fame.

4.4 Limit Repetition of the Same Top Artists in Playlists

Spotify should avoid showing the same top artists too frequently in algorithmic playlists by setting simple limits on repeated artist appearances.

Why this helps:
This improves playlist freshness and helps listeners discover a wider range of artists without reducing overall engagement.

4.5 Use Dashboards to Detect Rising Tracks Early

Spotify should develop dashboards that track early popularity trends and audio features to identify songs that are gaining attention quickly.

Why this helps:
Early identification allows Spotify to act faster by adding tracks to playlists or supporting them through promotion.

4.6 Provide Artists with Simple Performance Insights

Spotify should offer artists clear insights showing how their songs compare to genre averages in terms of energy, danceability, and emotional tone.

Why this helps:
This supports artists in making informed creative and marketing decisions and strengthens Spotify's relationship with creators.

5. Limitations

Despite its insights, this analysis has several limitations:

- The dataset does not include user-level listening behavior, preventing behavioral analysis.
- Popularity is treated as a static score, without accounting for short-term viral effects.
- Genre classification may oversimplify hybrid or evolving music styles.
- Correlation analysis does not imply causation between audio features and popularity.
- The time range is limited to 2015–2020 and may not reflect recent market changes.

6. Future Improvements

To enhance the analytical value of this BI solution, future work could include:

- Integrating user interaction data such as streams, skips, and playlist additions.
- Adding geographic dimensions to analyze regional preferences.
- Applying advanced analytics techniques such as regression or clustering.
- Incorporating real-time or near-real-time data updates.
- Expanding the model to support predictive popularity scoring for new releases.

7. Conclusion

This Business Intelligence project demonstrates how structured data modeling and interactive dashboards can transform raw music data into meaningful insights. The analysis highlights the importance of audio features, genre dynamics, and artist influence in shaping song popularity. By leveraging these insights, Spotify can refine its content strategies, enhance user experience, and maintain its competitive advantage in the evolving music streaming landscape.