

Spotify Music Streaming Analysis

Ahrar Karim, Kara Liao, Rita Feng, Kaixin Gao, Yanlun Li

Outline

- ★ **Problem Statement & Motivation**
- ★ **Data Overview**
- ★ **ER Diagram**
- ★ **EDA**
- ★ **Dashboards**
- ★ **Conclusion**
- ★ **Limitations**
- ★ **Recommendations**



01

***Problem Statement
&
Motivation***



Problem Statement

This project addresses the uncertainty around what drives streaming success on Spotify by analyzing how audio features, engagement metrics, and chart performance relate to track popularity across regions and over time. We explore which musical attributes and trends influence popularity, geographic variation, and chart longevity.

Motivation

Understanding what drives music popularity helps artists, producers, and Spotify teams better predict audience preferences and optimize content strategy. With streaming dominating today's music industry, data-driven insights into audio features, engagement metrics, and evolving listener trends can reveal what resonates with audiences and how tastes change over time.



02

***Data
Overview***



Data Overview

A comprehensive dump of Spotify’s Top 200 charts and approximately 0.9 million Spotify tracks from 2017 and continuing through the dataset’s last update

Field name	Type
track_id	STRING
date	DATE
region	STRING
chart	STRING
streams	INTEGER
track_name_canon	STRING
artist_name_norm	STRING
popularity	FLOAT
danceability	FLOAT
energy	FLOAT

valence	FLOAT
tempo	FLOAT
acousticness	FLOAT
liveness	FLOAT
instrumentalness	FLOAT
loudness	FLOAT
album_name	STRING
album_release_date	STRING
artist_popularity	FLOAT

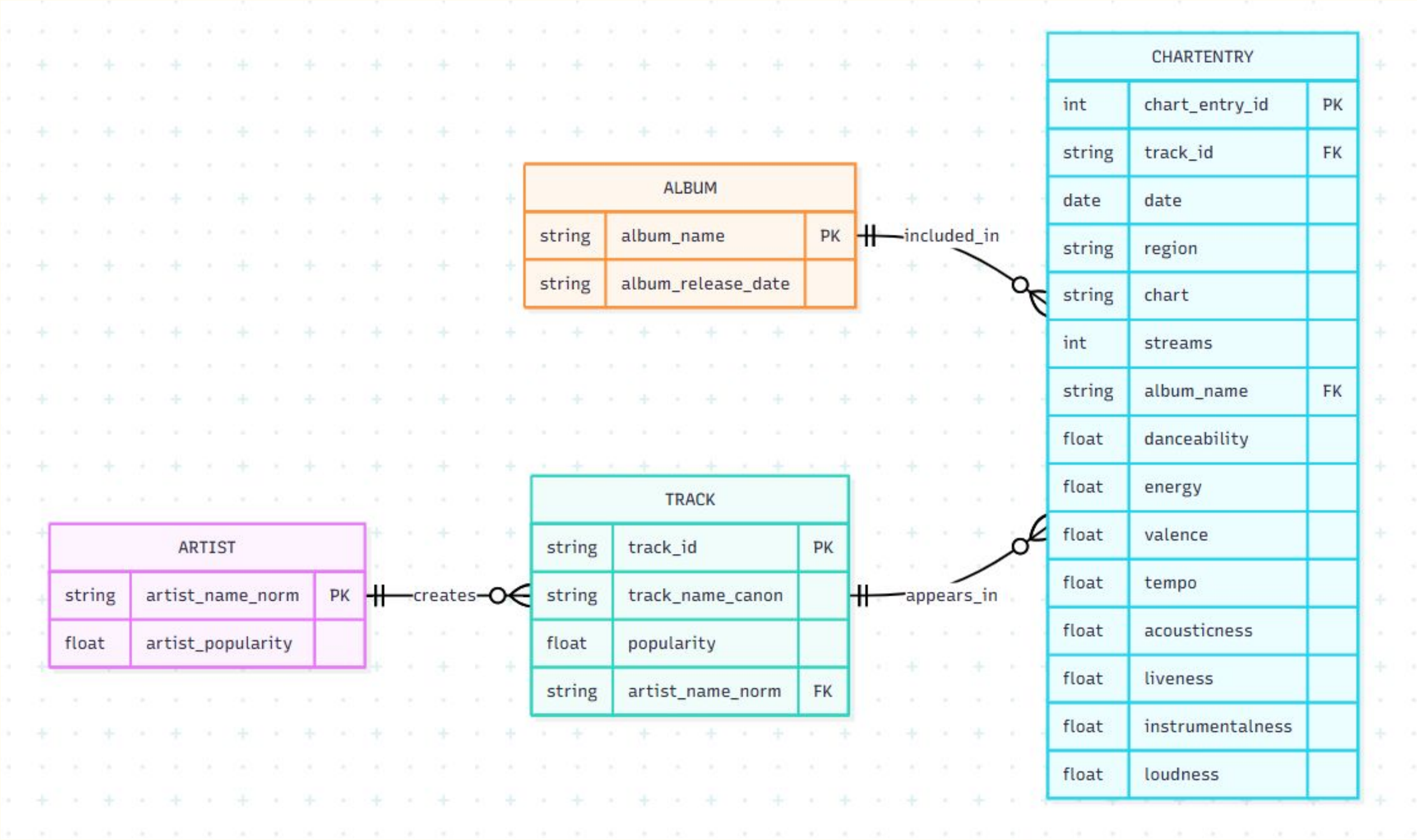


03

ER Diagram



ER Diagram



04

EDA



Impact of Artist Fame vs. Song Quality on Streaming Success

Region	Artist Fame Influence
Philippines	0.211
New Zealand	0.198
Argentina	0.186
Indonesia	0.166
Austria	0.165
Morocco	-0.028
Germany	-0.031

*Morocco and Germany represent the two lowest values.

Artist Fame Influence =
(Correlation between Artist Fame and Streams) –
(Correlation between Song Quality and Streams)

Positive value → Fame matters more than song qualities
Negative value → Song qualities matter more

Overall, fame has a stronger impact on streams in many regions, while a few regions lean more toward song-driven preferences.



Audio Feature Difference Across Popularity Level

Popularity Group	Danceability	Energy	Acousticness	Instrumentalness	Loudness
High	0.622	0.595	0.509	0.084	-7.9
Medium	0.697	0.649	0.532	0.009	-6.13
Low	0.689	0.648	0.511	0.018	-6.39

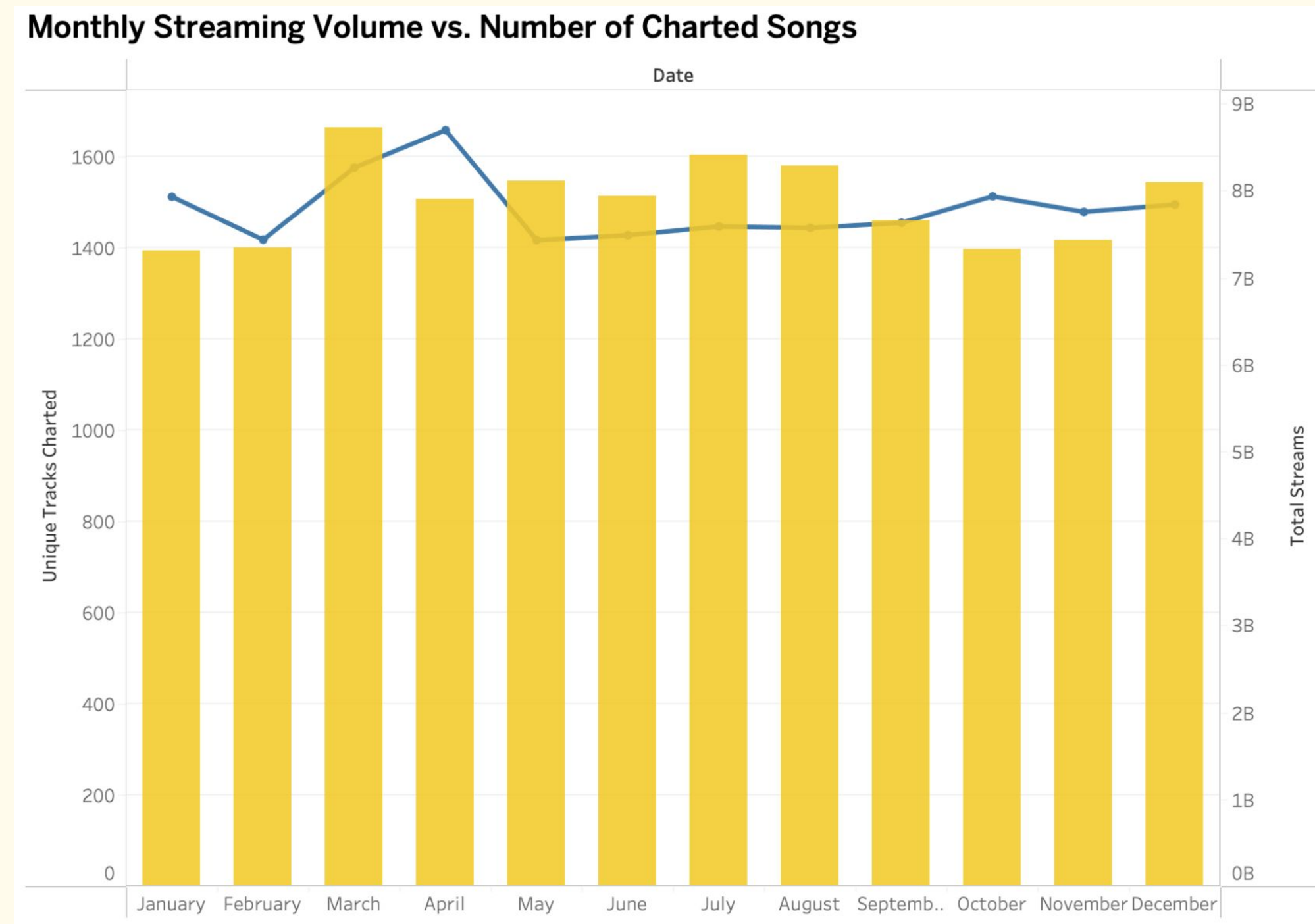
Medium-popularity songs have the highest **danceability and energy**, suggesting that more upbeat and engaging tracks tend to attract steady listener interest.

High-popularity songs stand out with higher **acousticness and instrumentalness**, indicating that distinct production styles, not just energy.

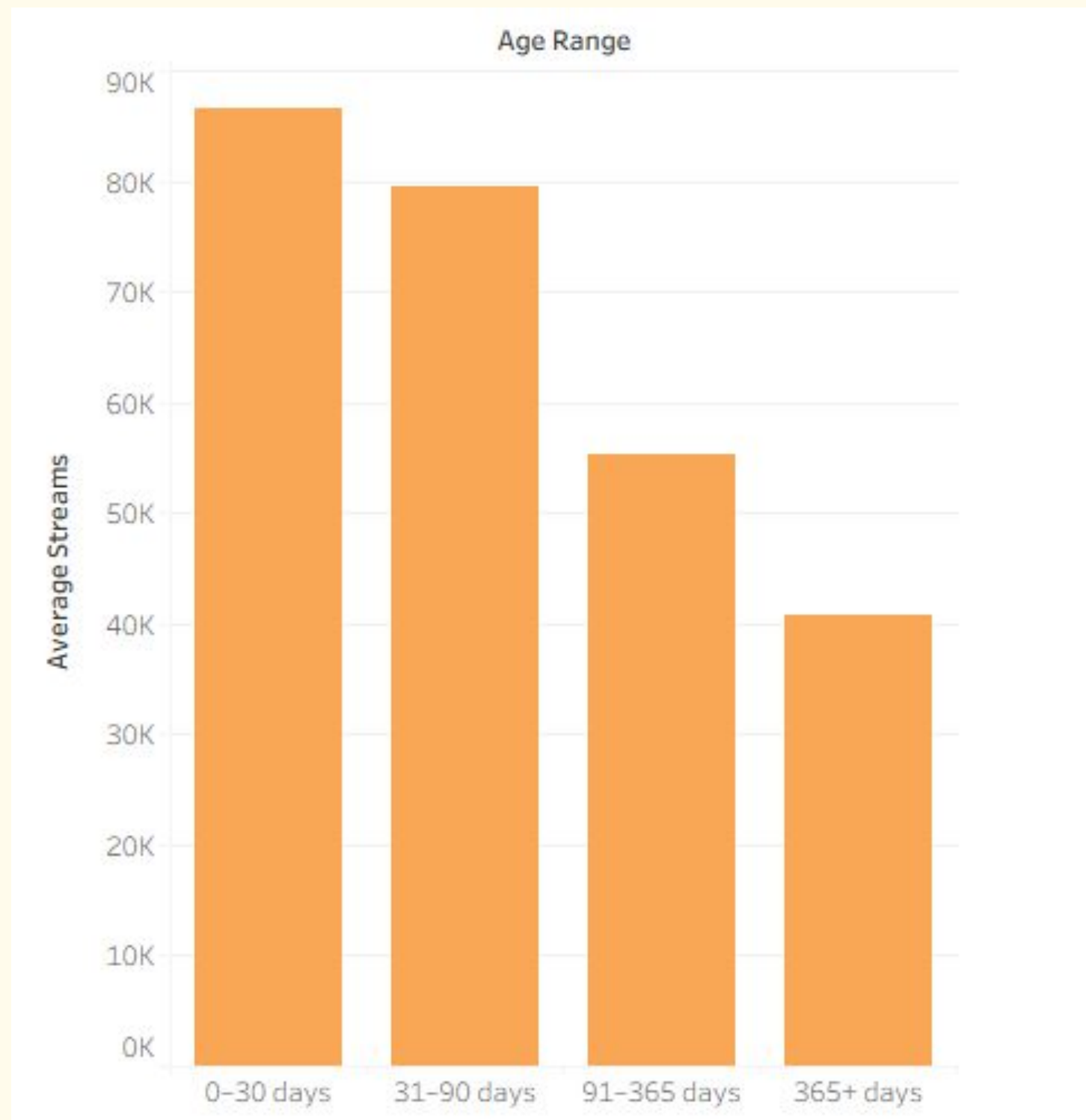


Monthly Streaming Patterns and Chart Activity

- ★ Chart activity **peaks from March to August** with more unique songs entering
- ★ Total streaming volume **stays relatively stable year-round**
→ consistent listener demand



Impact of Song Age on Streaming Performance



Newly released songs (0–30 days) receive the highest average streams, showing a clear recency effect where listeners favor fresh releases.

Streaming volume declines as songs age, indicating older tracks lose momentum over time and tend to have shorter chart longevity.



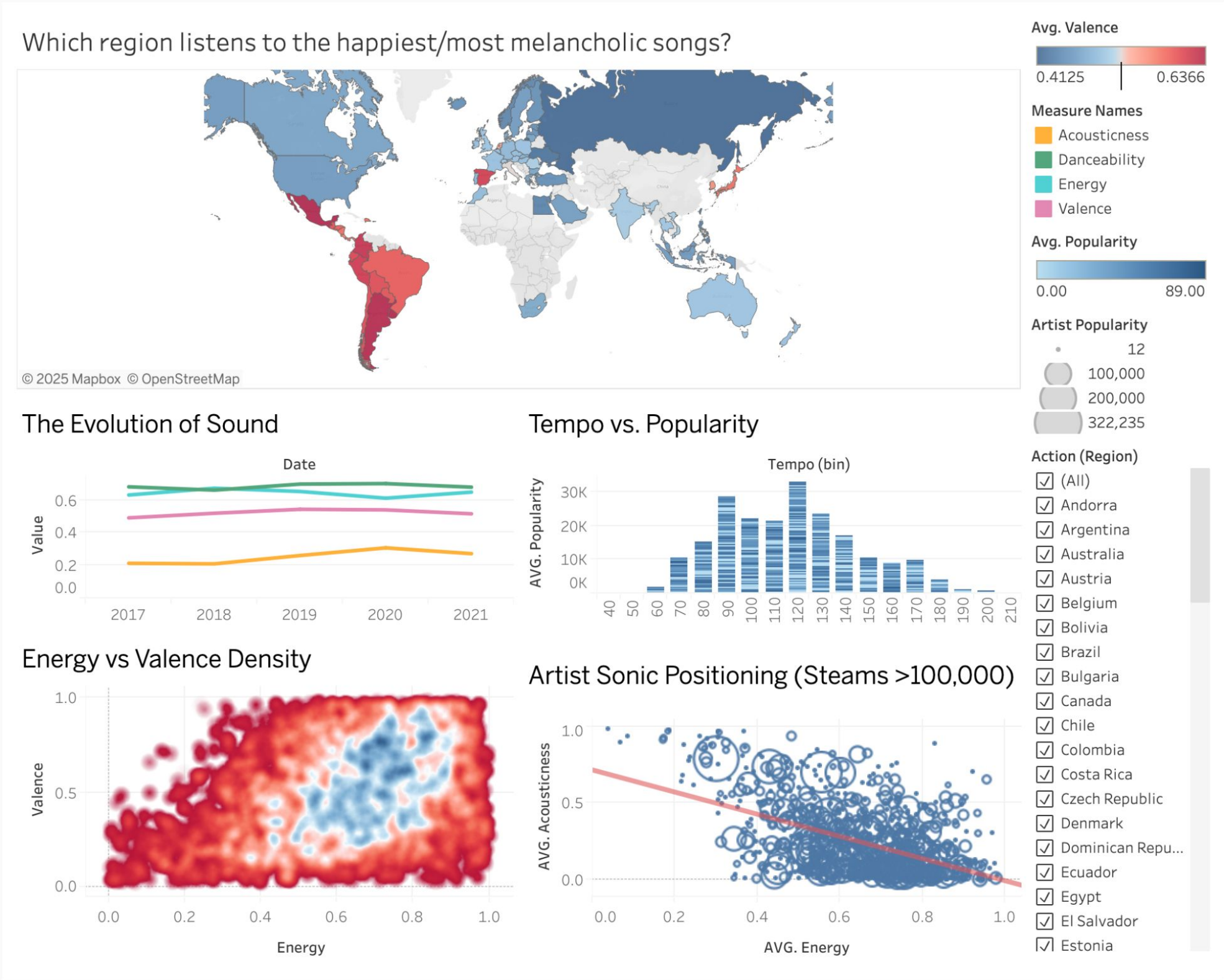
05

Dashboards



Dashboard

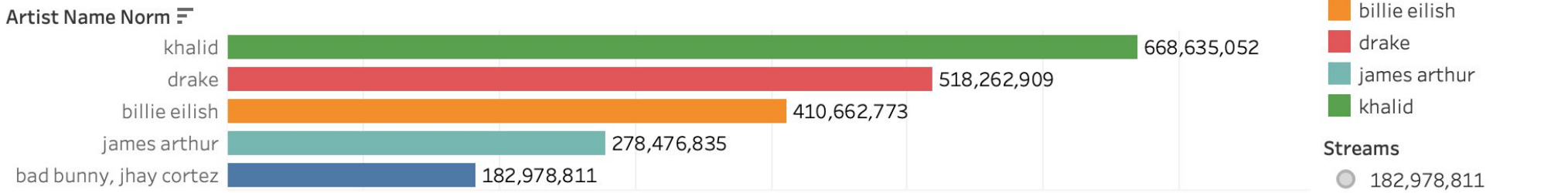
[Link](#)



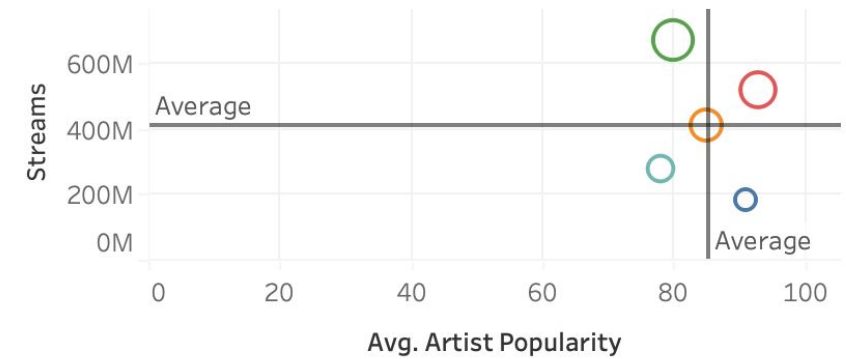
Dashboard

[Link](#)

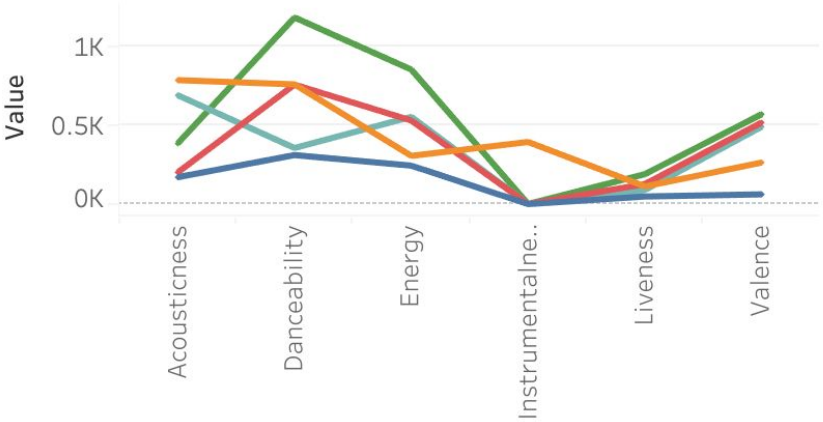
Top 5 Streaming Artists – US



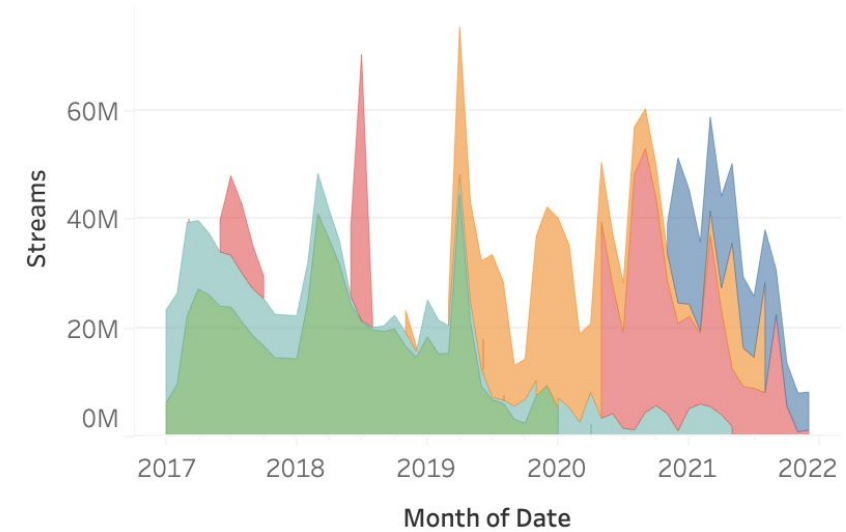
Top 5 Artists Positioning: Global Fame vs. US Streams



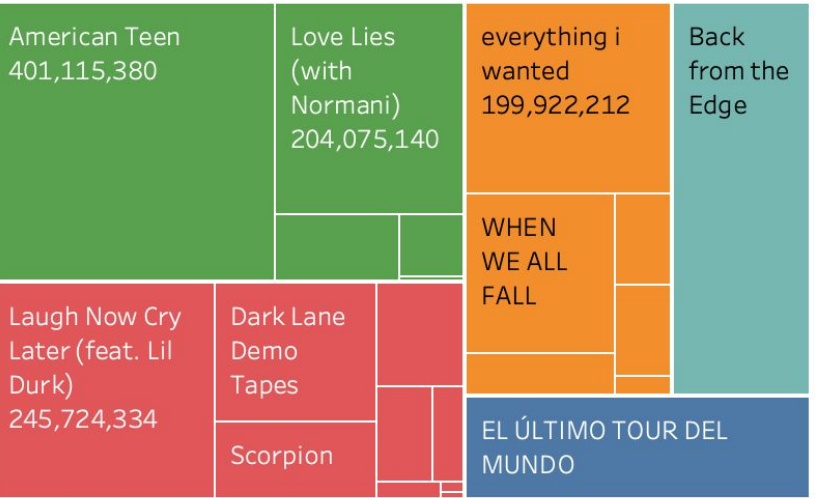
Audio Feature Profile (Top 5 Artists) - US



Monthly Streaming Trends (Top 5 Artists): 2017 - 2021 (US Market)



The Hit List: Which Album Drive the Success? (US Market)



06

Conclusion



Conclusion

- ★ Neutral-mood tracks with smoother, vocal-driven production consistently outperform songs with extreme emotions or aggressive loudness
- ★ Mid-tempo tracks (110–125 BPM) with balanced energy gain both strong global appeal and superior long-term chart longevity
- ★ Popularity tiers are highly stable over time—Medium stays Medium, Low stays Low—and true breakout transitions are extremely rare
- ★ Seasonality influences feature effectiveness: valence consistently peaks in summer, while danceability peaks in early spring
- ★ Artist branding can outweigh track features in markets like the Philippines or New Zealand, while in Poland or Korea, track-level characteristics matter more
- ★ Releasing more songs does not significantly increase popularity (correlation ≈ 0.13)



07

Limitations



Limitations

1

No genre-level analysis, so feature patterns across pop, hip-hop, EDM, or Latin may behave differently and could be masked by aggregation.

2

Regional data density and static popularity scores introduce bias, meaning some markets or lifecycle transitions may be over or under-represented.

3

External drivers like promotion, playlist placement, or social-media trends are not included, so observed relationships are correlational rather than causal.



08

Recommendations



Recommendations

-  Prioritize neutral-mood, mid-tempo, balanced-energy tracks for global promotion, as these features consistently deliver both broad appeal and strong longevity
-  Invest in strong brand-power artists rather than increasing release frequency, since branding drives performance far more than output volume
-  Adopt localized strategies based on regional feature preferences, such as high-energy rhythmic tracks in Latin markets and feature-driven positioning in track-centric regions
-  Use long-longevity tracks as playlist anchors, supplementing them with short-term high-danceability songs to balance stability and freshness



Thank You!

