Spotify Top 10,000 Songs from 1960-Now: Project Write-Up

Team 21: Michelle Saikali, Rhett Sessions, Ruiqi Cheng, Anthony Ayala, Chris Shen

Introduction:

In the dynamic music industry, where artists, genres, and audience preferences intertwine, data-driven insights have become fundamental. This project leverages data analytics to unveil patterns in artist-genre associations, user preferences, and track statistics. Exploring a diverse dataset of artists, albums, tracks, and genres, our goal is to gain strategic insights on the evolution of the music industry.

Business Problem:

The music industry has undergone a change in the digital era, with the rise of digital streaming platforms, such as Spotify and Apple Music. Artists and labels now have to navigate this landscape on understanding audience preferences, and predicting trends. To address these challenges, we aim to tackle key issues:

Genre-Artist Relationship: Identify and quantify relationships between artists and genres.

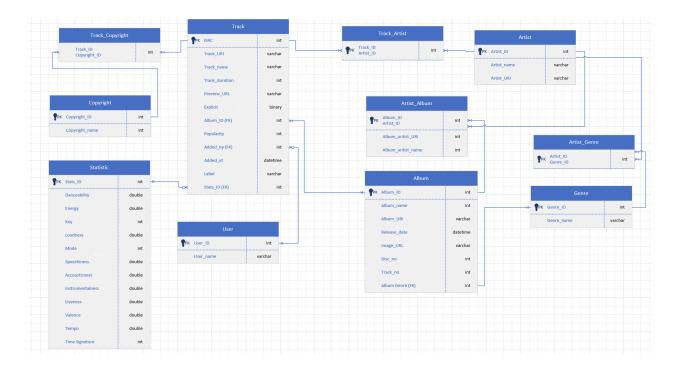
User Engagement: Analyze user preferences and engagement patterns. Track Performance: Evaluate individual track performance based on various statistics.

Album-Artist Associations: Explore associations between albums and artists.

By addressing these issues, we have gained actionable insights on the music industry and evolution of time.

Table Design/Relationships:

- Track Copyright: many-to-many
- Track Statistic: one-to-many
- Track User: one-to-many
- Track Artist: many-to-many
- Track Album: one-to-many
- Artist Genre: many-to-many
- Artist Album: many-to-many
- Album Genre: one-to-many



Observations and Results:

(1) What are the top 5 popular tracks and (2) their artists in the subsetted database? This in addition to research about the songs gave us insights on what types of songs were the most popular from our dataset. It seems that rock songs tend to be the most popular due to their longevity.

track_name		track_uri		popularity
Here Without You		spotify:track:3NLrRZoMF0Lx6zTlYqeIo4		80
Dont Stop - 2004 Remaster		spotify:track:4bEb3KE4mSKITFjtWJQBqO		79
Eastside (with Halsey & Khalid)		spotify:track:0d2iYfpKoM0QCKvcLCkBao		78
Juke Box Hero		spotify:track:00qOE7OjRl0BpYiCiweZB2		74
Its Like That		spotify:track:0CtkjgZpkgnW7U6WmHsakD		68
popularity	Artist_URI		Artist_Name	
80	spotify:artist:2RTUTCvo6onsAnheUk3aL9		3 Doors Down	
79	spotify:artist:08GQAI4eElDnROBrJRGE0X		Fleetwood Mac	
78	spotify:artist:5CiGnKThu5ctn9pBxv7DGa		benny blanco, Halsey, Khalid	
74	spotify:artist:6IRouO5mvvfcyxtPDKMYFN		Foreigner	
64	spotify:artist:0TnOYISbd1XYRBk9myaseg		Pitbull	

(3) We hypothesized that the most upbeat songs would be the most popular, but there was only one song whose danceability, energy, and liveness were above the average, and its popularity rating is zero. So it turns out that these three songs don't seem to be significant in determining popularity. (4) However, shorter songs tend to have much higher popularity rankings on average. Over the years, artists have shortened the length of their songs so that they

track_name	popularity					
Justified & Ancient - Stand by the Jams 0						
track_duration_category	average_popularity	/				
more than 3 minutes	47.6364					
less than 3 minutes	78.0000					
	.,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,					

can generate more listens, so it makes sense that shorter songs are more popular.

(5) Finally, we determined the least labels with the least popular tracks. Interestingly, some of these songs are recognizable, so it makes us question the validity of the popularity ratings.

label	max(popularity)	track_name
Jams Communications	0	Justified & Ancient - Stand by the Jams
Universal Music Group	0	Mercy
Universal Music Group	0	You Cant Always Get What You Want
Sanctuary Records	42	Apeman - 2014 Remastered Version
Jive	56	From the Bottom of My Broken Heart