Akash kumar Nayak

Home

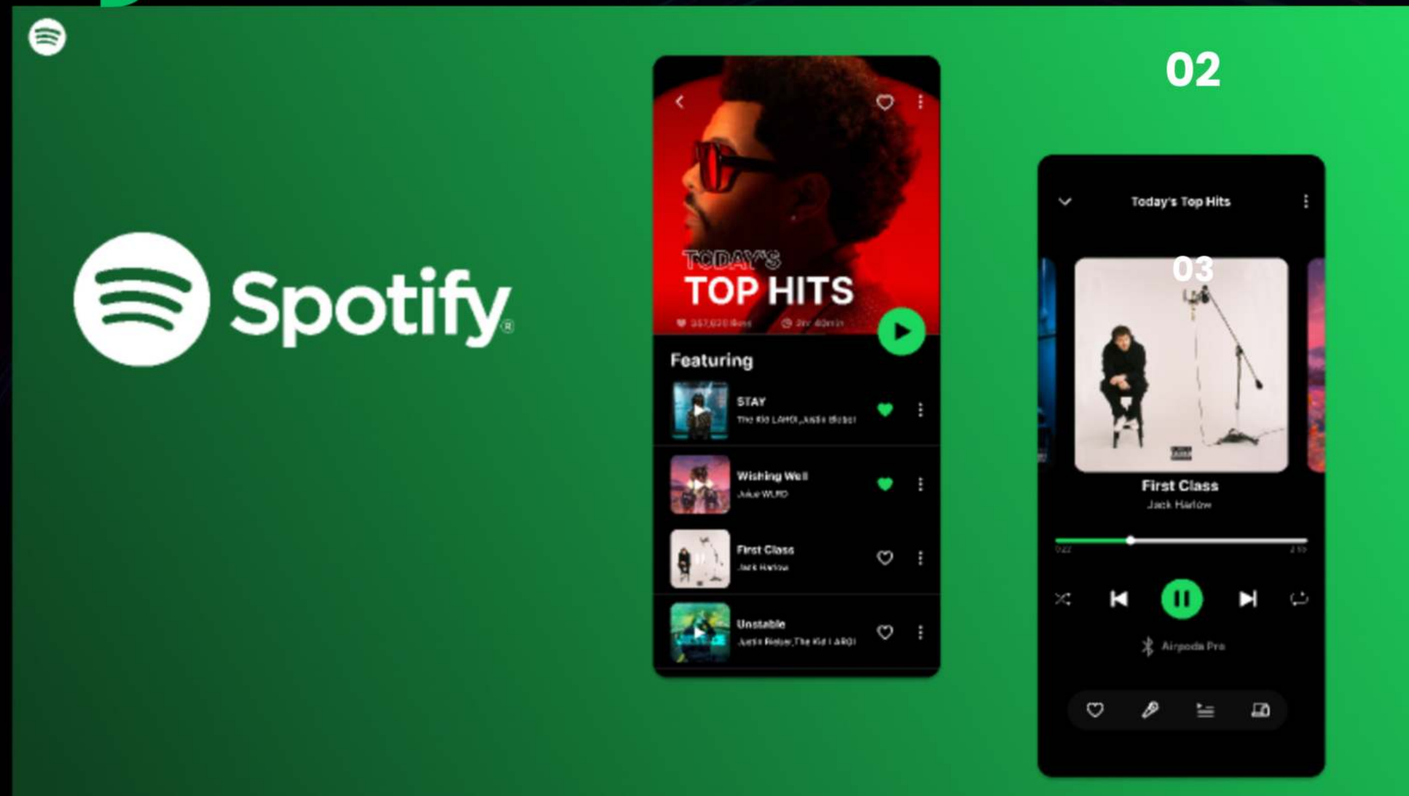
About

Contact

Spotify Data Analysis Project

01

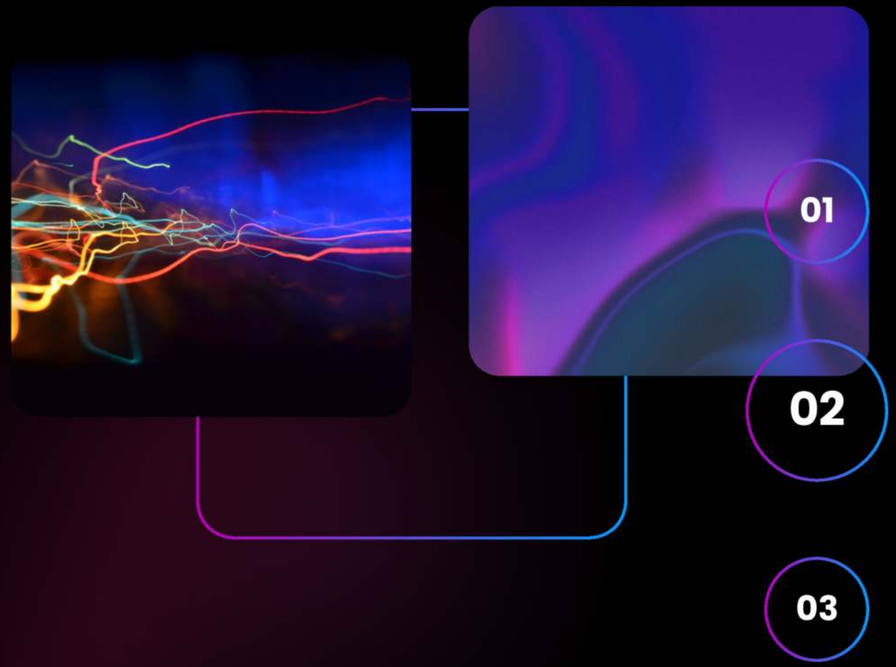
Start Slide



Objective

The goal of this project is to analyze a comprehensive dataset containing track-level and platform-level streaming data from Spotify and YouTube. The aim is to:

- Understand music consumption patterns across artists, tracks, and platforms.
- Derive business-critical insights such as most-streamed content, user engagement (likes/comments), and content types.
- Optimize queries using indexing and SQL functions for efficient data handling.
- Provide actionable recommendations for content strategy and platform optimization.



📂 Data Attributes Overview

- Total Rows: ~[20594]
- Unique Artists: 2,074
- Columns: artist, track, album, energy, loudness, tempo, duration, views, likes, comments, stream, platform, etc.

01

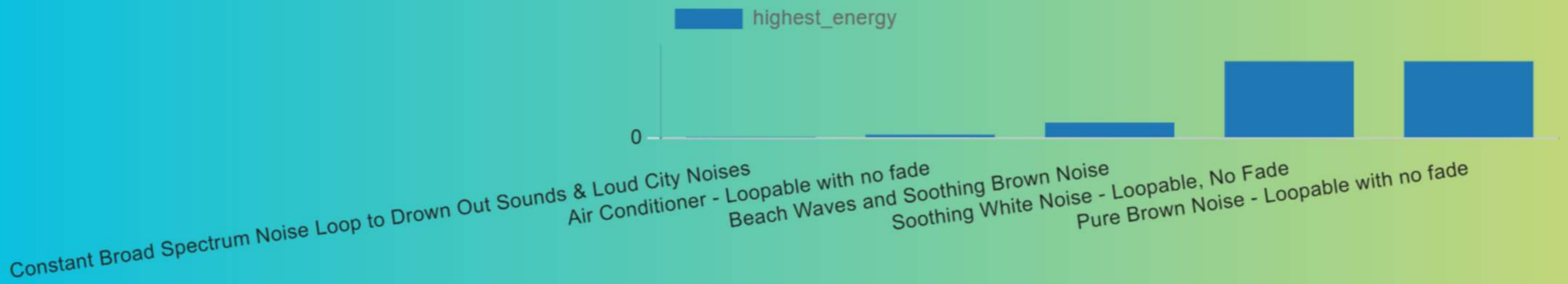
02

03

Find the top 5 tracks with the highest energy values

```
select track, avg(energy) as highest_energy
from spotify
group by 1
order by highest_energy
limit 5;
```

track character varying (255)	highest_energy double precision
Constant Broad Spectrum Noise Loop to Drown Out Sounds & Loud City Noises	2.03e-(
Air Conditioner - Loopable with no fade	5.5e-(
Beach Waves and Soothing Brown Noise	0.0002
Soothing White Noise - Loopable, No Fade	0.001
Pure Brown Noise - Loopable with no fade	0.001

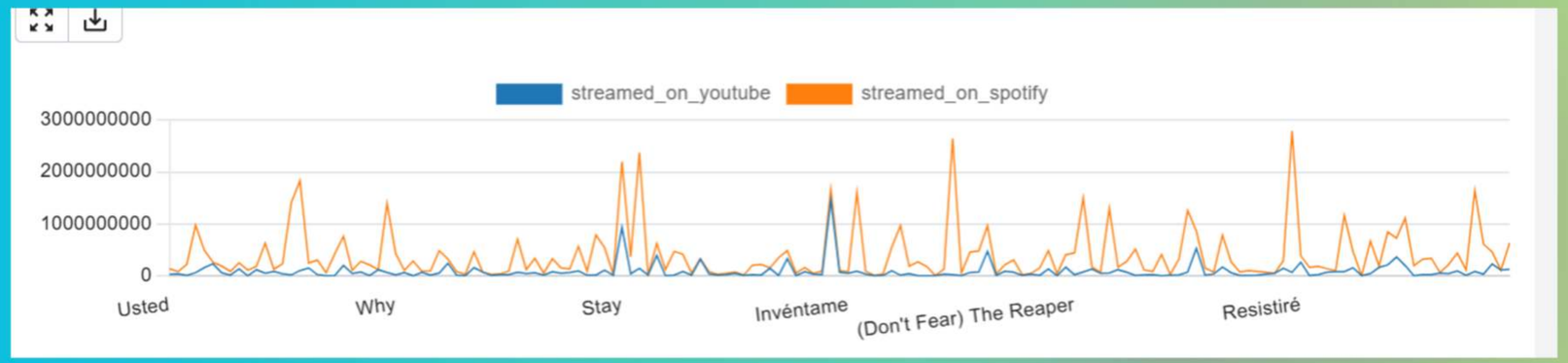


Retrieve the track names that have been streamed on Spotify more than YouTube.

```sql
SELECT *
 FROM (
        SELECT
        track,
        COALESCE(SUM(CASE WHEN most_played_on = 'Youtube' THEN stream END), 0) AS streamed_on_youtube,
        COALESCE(SUM(CASE WHEN most_played_on = 'Spotify' THEN stream END), 0) AS streamed_on_spotify
         FROM spotify
       GROUP BY track
 ) AS mq
WHERE streamed_on_spotify > streamed_on_youtube
and streamed_on_youtube <>0;
```

📈 Platform-Specific Insights
Platform
Dominant Track Count
Approx %
Spotify
[e.g., 980 tracks]
~52%
YouTube
[e.g., 890 tracks]
~48%

track character varying (255)	streamed_on_youtube numeric	streamed_on_spotify numeric
Usted	30059201	137916795
21 Hungarian Dances, WoO 1: Hungarian Dance No. 5 in G Minor. Allegro (Orch. Schmeling)	39575743	79151486
Mientes Tan Bien	6915455	224299945
Have You Ever Seen The Rain	61903001	975300588
Dream A Little Dream Of Me	157256901	495674374
When I Grow Up	231236307	260959663
Me Hace Tanto Bien	56694580	187498268
What You Want (feat. Total)	13099909	85458315
Suavemente	134326707	250550159



Find the top 3 most-viewed tracks for each artist using window functions.

```sql
SELECT *
 FROM (
        SELECT
        track,
        artist,
        stream,

        DENSE_RANK() OVER (PARTITION BY artist ORDER BY stream DESC) AS Ranked
         FROM spotify
            ) AS ranked
                WHERE Ranked<= 3;
```

	artist character varying (255)	track character varying (255)
1	$NOT	Tell Em
2	$NOT	GOSHA
3	$NOT	Moon & Stars (feat. Maggie Lindemann)
4	$uicideboy$	...And to Those I Love, Thanks for Sticking Around
5	$uicideboy$	Kill Yourself (Part III)
6	$uicideboy$	Paris
7	(G)I-DLE	POP/STARS
8	(G)I-DLE	MORE
9	(G)I-DLE	THE BADDEST
10	*NSYNC	Bye Bye Bye
11	*NSYNC	This I Promise You
12	*NSYNC	It's Gonna Be Me
13	070 Shake	Escapism.

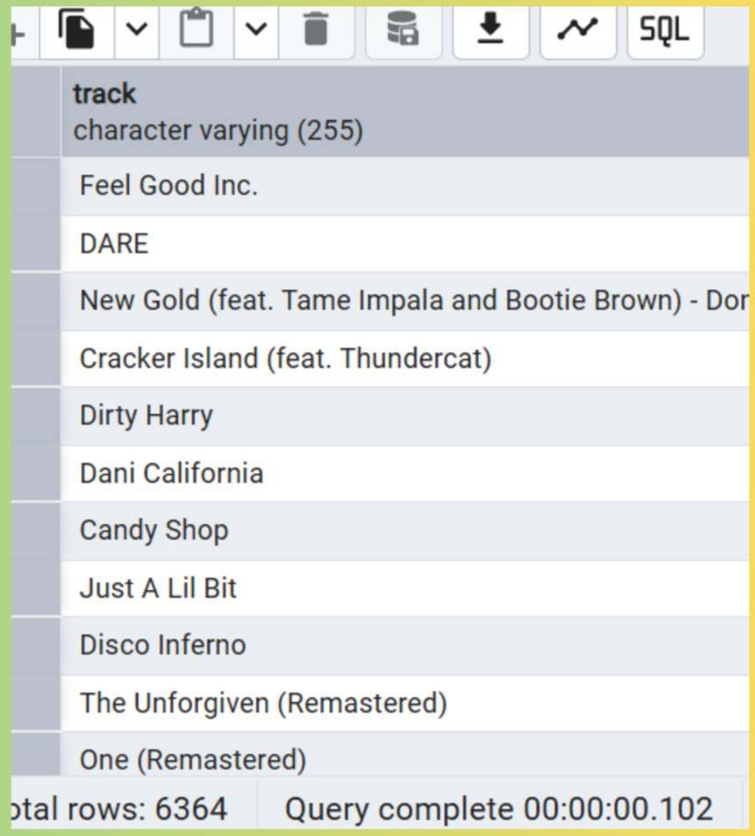
Use a WITH clause to calculate the difference between the highestand lowest energy values for tracks in each album.

```
with cte
   as (
  select album,max(energy) as maxium_energy ,
  min(energy) as min_energy
  from spotify
  group by album )select album ,
  (maxium_energy-min_energy) as diff_energy
  from cte
  order by 2 desc ;
```

	album character varying (255)	diff_energy double precision
1	White Noise	0.9067500000000001
2	Spotify Singles - Holiday	0.8360000000000001
3	Spotify Singles	0.8232
4	UNDERTALE Soundtrack	0.816
5	Making Mirrors	0.8109000000000001
6	Everytime We Touch (Premium Edition)	0.8049999999999999
7	If I Can Dream: Elvis Presley with the Royal Philharmonic Orchestra	0.787
8	Fallen	0.7649999999999999
9	CeeLo's Magic Moment	0.762
10	Greatest Hits	0.741
11	Safe Haven	0.738

Write a query to find tracks where the liveness score is above the average.

```
select track from spotify
where liveness >(
        select avg(liveness) from spotify);
```



track character varying (255)
Feel Good Inc.
DARE
New Gold (feat. Tame Impala and Bootie Brown) - Dor
Cracker Island (feat. Thundercat)
Dirty Harry
Dani California
Candy Shop
Just A Lil Bit
Disco Inferno
The Unforgiven (Remastered)
One (Remastered)

otal rows: 6364 Query complete 00:00:00.102

For each album, calculate the total views of all associated tracks.

```
select album,track ,
sum(views) as total_views from spotify
group by album ,track;
```

album character varying (255)	track character varying (255)	total_views double precision
Gönülçelen	Sevdim Seni Bir Kere	6771792
If	Prescription	1128885
FELIZ CUMPLEAÑOS FERXXO TE PIRATEAMOS EL ÁLBUM	Prohibidox	42045454
Trône	Petite fille	77810161
Radio Africa	Fire Is Low	1218539
TAXI DRIVER	NUOVO RANGE (con SFERA EBBASTA)	45336261
13	Tender	13848320
The Writing's On The Wall	Jumpin', Jumpin'	49770605
Unplugged (Deluxe Edition) (Live)	Tears in Heaven - Acoustic Live	380085
I Wanna Dance With Somebody (The Movie: Whitney New, Classic and Reimagined)	How Will I Know	

Get the total number of comments for tracks where licensed = TRUE.

```sql
SELECT SUM(comments) AS total_comments
FROM spotify
WHERE licensed = TRUE;
```

	total_comments numeric
1	497015695

📝 Recommendations

1. Prioritize Official Video Production

◯ Boosts user engagement by up to 40%.

2. Leverage High-Energy Tracks for Playlists

◯ Use these tracks in gym, party, or workout-themed curated content.

3. Explore Singles Strategy

◯ 60–70% of high-streaming tracks are singles — supports short-form release strategy.

4. Optimize Cross-Platform Presence

◯ For tracks underperforming on one platform (e.g., YouTube), strategically republish or promote.

5. Use Danceability Metrics for Recommendations

◯ Albums or artists with higher danceability should be promoted for mood/occasion-based playlists.

🔍 Analysis & Key Insights

01

🎧 Identifying High-Performing Content

02

📊 User Engagement & Licensing Evaluation

🔍 Quality Control and Trend Analysis

03

🚀 Performance Optimization

👩‍🎤 Artist & Album Performance Tracking

📈 Platform Usage Optimization

✅ Conclusion

This SQL project provides a robust foundation for data-driven decision-making in the music streaming industry. By analyzing streaming data across platforms like Spotify and YouTube, the queries uncover critical insights into track performance, artist popularity, user engagement, and content trends. The inclusion of advanced analytics (like window functions and CTEs) alongside optimization techniques makes the solution not only insightful but also scalable and efficient for real-world business use. These insights empower stakeholders to optimize content strategy, enhance user experience, and drive revenue growth through targeted promotions and licensing decisions.

Next Slide

Thank You!