Data Analysis Queries
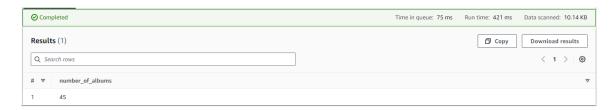
Q1: a) Number of Albums

```
SELECT count(distinct album_id) as number_of_albums
FROM "spotify_database"."album_data";
```

⊘ Completed	Time in queue: 75 ms	Run time: 421 ms	Data scanned: 10.14 KB

Results (1)   Copy   Download results

🔍 Search rows   ‹ 1 › ⚙

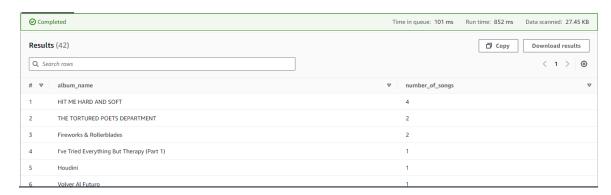| # ▽	number_of_albums
1	45

b) Number of Albums having clean release dates

```
SELECT count(distinct album_id) as count_of_albums_with_dates
FROM "spotify_database"."album_data"
where ablum_release_date is not null;
```

⊘ Completed	Time in queue: 83 ms	Run time: 464 ms	Data scanned: 10.14 KB

Results (1)   Copy   Download results

🔍 Search rows   ‹ 1 › ⚙

# ▽	count_of_albums_with_dates
1	43

Q2: a) Number of songs from each album

```
SELECT al_d.album_name,
        count(distinct s_d.song_id) as number_of_songs
FROM "spotify_database"."album_data" al_d
        inner join "spotify_database"."songs_data" as s_d on al_d.album_id = s_d.album_id
where al_d.ablum_release_date is not null
group by al_d.album_name
order by number_of_songs desc;
```

⊘ Completed	Time in queue: 101 ms	Run time: 852 ms	Data scanned: 27.45 KB

Results (42)   Copy   Download results

🔍 Search rows   ‹ 1 › ⚙

# ▽	album_name ▽	number_of_songs ▽
1	HIT ME HARD AND SOFT	4
2	THE TORTURED POETS DEPARTMENT	2
3	Fireworks & Rollerblades	2
4	I've Tried Everything But Therapy (Part 1)	1
5	Houdini	1
6	Volver Al Futuro	1

b) Most successful albums – (albums having more than one song in the top played song's list)?

```
WITH A as (
        SELECT al_d.album_name,
```

```
                count(distinct s_d.song_id) as count_album_id
        FROM "spotify_database"."album_data" al_d
                inner join "spotify_database"."songs_data" as s_d on s_d.album_id = al_d.album_id
        where al_d.ablum_release_date is not null
        group by al_d.album_name
        order by count_album_id desc
)
Select A.album_name,
        A.count_album_id
FROM A
WHERE A.count_album_id > 1;
```

⊘ Completed			Time in queue: 95 ms	Run time: 807 ms	Data scanned: 27.45 KB

Results (3)   Copy   Download results

# ▽	album_name ▽	count_album_id ▽
1	HIT ME HARD AND SOFT	4
2	THE TORTURED POETS DEPARTMENT	2
3	Fireworks & Rollerblades	2

Q3: a) Most popular artists – (artists having more than one song in the top played song's list)?

```
    WITH A as (
        SELECT ar_d.artist_name,
                count(distinct s_d.song_id) as count_artist_id
        FROM "spotify_database"."artist_data" ar_d
                inner join "spotify_database"."songs_data" as s_d on ar_d.artist_id = s_d.artist_id
                inner join "spotify_database"."album_data" as al_d on s_d.album_id = al_d.album_id
        where al_d.ablum_release_date is not null
        group by ar_d.artist_name
        order by count_artist_id desc
)
Select A.artist_name,
        A.count_artist_id
FROM A
WHERE A.count_artist_id > 1;
```
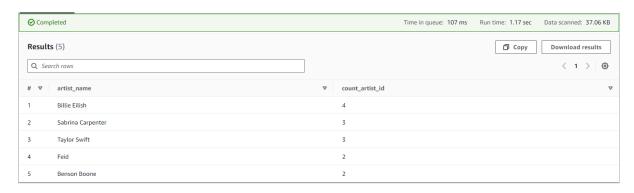
⊘ Completed			Time in queue: 107 ms	Run time: 1.17 sec	Data scanned: 37.06 KB

Results (5)   Copy   Download results

# ▽	artist_name ▽	count_artist_id ▽
1	Billie Eilish	4
2	Sabrina Carpenter	3
3	Taylor Swift	3
4	Feid	2
5	Benson Boone	2

b) Number of most popular artists – (artists having more than one songs in the top played song's list)

```
WITH A as (
        SELECT s_d.artist_id,
                count(distinct s_d.song_id) as count_artist_id
        FROM "spotify_database"."artist_data" ar_d
                inner join "spotify_database"."songs_data" as s_d on ar_d.artist_id = s_d.artist_id
                inner join "spotify_database"."album_data" as al_d on s_d.album_id = al_d.album_id
        where al_d.ablum_release_date is not null
        group by s_d.artist_id
        order by count_artist_id desc
)
Select count(distinct A.artist_id) as number_of_popular_artists
FROM A
WHERE A.count_artist_id > 1;
```
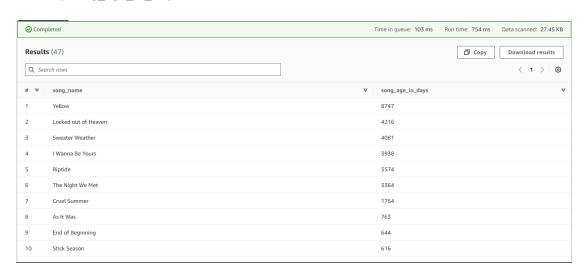
⊘ Completed	Time in queue: 75 ms	Run time: 910 ms	Data scanned: 37.06 KB

Results (1)   Copy   Download results

Q Search rows   ‹ 1 › ⚙

# ▽	number_of_popular_artists	▽
1	5	

Q4: a) Age of songs – (differences in Dates of Release and Date of Added in songs list)

```
SELECT distinct s_d.song_name,
        DATE_DIFF('day', al_d.ablum_release_date, s_d.song_added) as song_age_in_days
FROM "spotify_database"."album_data" al_d
        inner join "spotify_database"."songs_data" as s_d on al_d.album_id = s_d.album_id
where al_d.ablum_release_date is not null
order by song_age_in_days desc;
```

⊘ Completed	Time in queue: 103 ms	Run time: 754 ms	Data scanned: 27.45 KB

Results (47)   Copy   Download results

Q Search rows   ‹ 1 › ⚙

# ▽	song_name	▽	song_age_in_days	▽
1	Yellow		8747	
2	Locked out of Heaven		4216	
3	Sweater Weather		4081	
4	I Wanna Be Yours		3938	
5	Riptide		3574	
6	The Night We Met		3364	
7	Cruel Summer		1764	
8	As It Was		763	
9	End of Beginning		644	
10	Stick Season		616	

b) Percentage of songs in each song age category – (Latest, Medium Old, Old)

```
With song_age_calculation as (
        SELECT distinct s_d.song_id,
                DATE_DIFF('day', al_d.ablum_release_date, s_d.song_added) as song_age_in_days
        FROM "spotify_database"."album_data" al_d
                inner join "spotify_database"."songs_data" as s_d on al_d.album_id = s_d.album_id
        where al_d.ablum_release_date is not null
        order by song_age_in_days desc
),
B as (
        Select song_age_calculation.song_id,
                song_age_calculation.song_age_in_days,
                CASE
                        WHEN song_age_in_days > 1095 THEN 'Very Old Song'
                        WHEN (
                                182 < song_age_in_days
                                AND song_age_in_days < 1095
                        ) THEN 'Medium Old Song' ELSE 'Latest_Song'
                END as song_age_category
        FROM song_age_calculation
)
select Song_Age_Category,
        count(distinct B.song_id) as Number_of_Songs_in_Category,
        CAST(
                (count(distinct B.song_id) * 100 /(select count(*)from song_age_calculation)) as real
        ) as Percent_of_Songs_in_Category
from B
group by song_age_category
order by song_age_category;
```

⊘ Completed		Time in queue: 102 ms	Run time: 939 ms	Data scanned: 54.90 KB

Results (3)   Copy   Download results

# ▽	Song_Age_Category ▽	Number_of_Songs_in_Category ▽	Percent_of_Songs_in_Category ▽
1	Latest_Song	32	68.0
2	Medium Old Song	8	17.0
3	Very Old Song	7	14.0

c) 5 songs in each age category

```
With song_age_calculation as (
        SELECT distinct s_d.song_name,
                s_d.song_id,
                DATE_DIFF('day', al_d.ablum_release_date, s_d.song_added) as song_age_in_days
        FROM "spotify_database"."album_data" al_d
                inner join "spotify_database"."songs_data" as s_d on al_d.album_id = s_d.album_id
```

```sql
        where al_d.ablum_release_date is not null
        order by song_age_in_days desc
),
B as (
        Select song_age_calculation.song_name,
                song_age_calculation.song_age_in_days,
                CASE
                        WHEN song_age_in_days > 1095 THEN 'Very Old Song'
                        WHEN (
                                182 < song_age_in_days
                                AND song_age_in_days < 1095
                        ) THEN 'Medium Old Song' ELSE 'Latest_Song'
                END as song_age_category
        FROM song_age_calculation
),
C as (
        select b.song_name,
                b.song_age_category,
                DENSE_RANK() OVER (
                        PARTITION BY song_age_category
                        ORDER BY song_age_in_days asc
                ) AS ranking
        from B
)
Select song_name,
        song_age_category,
        ranking
from C
where ranking <= 5
order by song_age_category asc;
```

Results (16)

Copy   Download results

Search rows

< 1 >

#	song_name	song_age_category	ranking
1	SORRY 4 THAT MUCH	Latest_Song	1
2	Please Please Please	Latest_Song	2
3	Houdini	Latest_Song	3
4	Volver Al Futuro	Latest_Song	4
5	Pink Skies	Latest_Song	5
6	LUNA	Medium Old Song	1
7	Lose Control	Medium Old Song	2
8	greedy	Medium Old Song	2
9	FE!N (feat. Playboi Carti)	Medium Old Song	3
10	Feather	Medium Old Song	4
11	Stick Season	Medium Old Song	5
12	Cruel Summer	Very Old Song	1
13	The Night We Met	Very Old Song	2
14	Riptide	Very Old Song	3
15	I Wanna Be Yours	Very Old Song	4
16	Sweater Weather	Very Old Song	5