SQL project: Spotify's Most Streamed Songs of 2023:

Dataset: https://www.kaggle.com/datasets/nelgiriyewithana/top-spotify-songs-2023

Question 1: Which artist has the most songs in the top streamed list of 2023?

Query:

```
select artist_name, count(*) as song_count from spotify_dataset.spotify_2023
where released_year = 2023
group by artist_name
order by song_count desc
limit 1
```

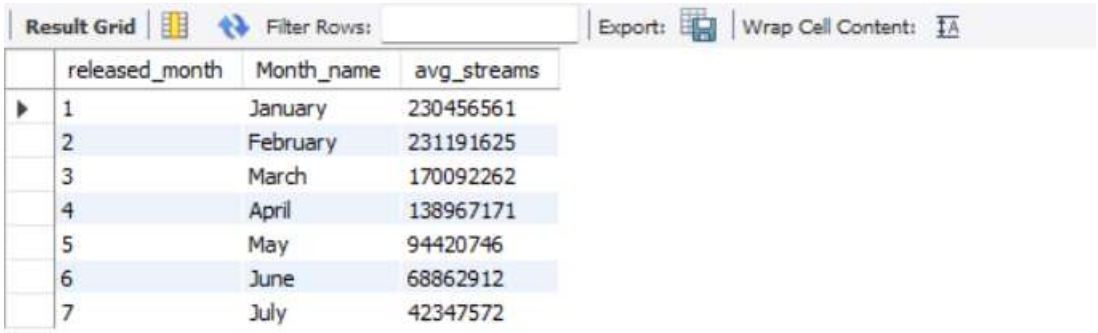
output:

Question 2: What is the average number of streams for songs released in each month of 2023?

Query:

```
select released_month,
case when released_month=1 then "January"
        when released_month=2 then "February"
        when released_month=3 then "March"
        when released_month=4 then "April"
        when released_month=5 then "May"
        when released_month=6 then "June"
        when released_month=7 then "July"
   end as Month_name,
round(avg(streams), 0) as avg_streams
from spotify_dataset.spotify_2023
where released_year = 2023
group by 1
order by 1
```

Output:



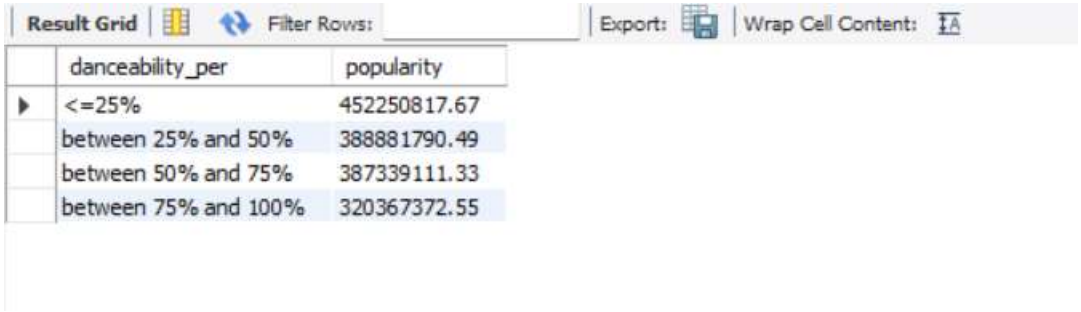
released_month	Month_name	avg_streams
1	January	230456561
2	February	231191625
3	March	170092262
4	April	138967171
5	May	94420746
6	June	68862912
7	July	42347572

Question 3: Are songs with higher danceability percentages generally more popular (i.e. have more streams)?

Query:

```
select
case
  when danceability_per >=0 and danceability_per<=25 then "<=25%"
  when danceability_per>25 and danceability_per<=50 then "between 25% and 50%"
  when danceability_per > 50 and danceability_per<=75 then "between 50% and 75%"
  when danceability_per >75 and danceability_per<=100 then "between 75% and 100%"
  end as danceability_per,
  round(avg(streams), 2) as popularity
  from spotify_dataset.spotify_2023
  group by 1
  order by 2 desc
```

Output:



danceability_per	popularity
<=25%	452250817.67
between 25% and 50%	388881790.49
between 50% and 75%	387339111.33
between 75% and 100%	320367372.55

Question 4 : What percentage of songs in the 'Top Spotify Songs 2023' were actually released in 2023?

Query:

```
WITH cte AS
(SELECT
COUNT(*) AS total_count,
SUM(CASE WHEN released_year = 2023 THEN 1 ELSE 0 END) AS count_2023
FROM spotify_dataset.spotify_2023)
select (count_2023*100)/total_count as Top_Spotify_Songs_2023
from cte
```

Output: