**TOPIC:**

**SQL Project**

**CREATED BY**: MATTA SAI PRAKASH

**CREATED DATE**: 02-02-2024.

**APPROVED BY**:

**Project:**

Spotify top songs analysis

1. Which artist has the most songs in the top streamed list of 2023?
2. What is the average number of streams for songs released in each month of 2023?
3. Are songs with higher danceability percentages generally more popular (i.e. have more streams)?
4. What percentage of songs in the 'Top Spotify Songs 2023' were released in 2023?

**Solution:**

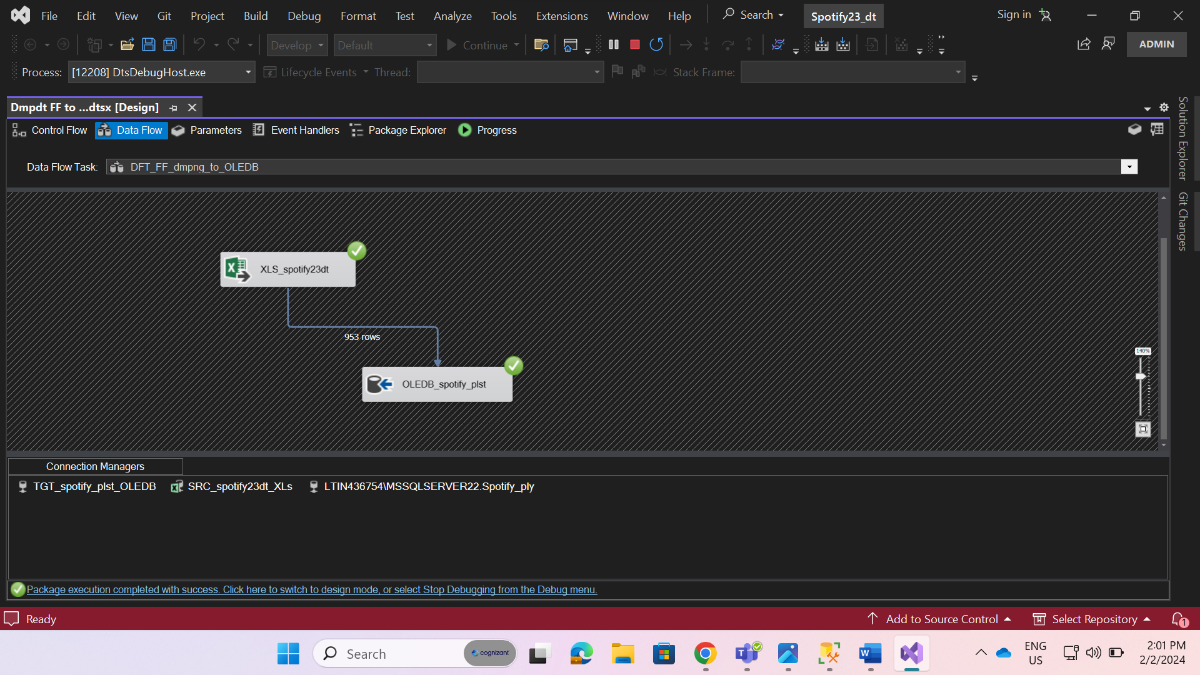
For Loading data to OLEDB (database) I have used SSIS (SQL SERVER INTEGRATION SERVICES)

**A screenshot of a computer

Description automatically generated**

Deleting old records using Execute SQL task

Data flow task used to set up for loading data from source file (XL) to target file (OLEDB)

****

Select \* from Spotify\_plst

Select top 10

[artist(s)\_name],

count(\*) as most\_played

from Spotify\_plst

where released\_year=2023

group by [artist(s)\_name]

order by most\_played desc

Select round(avg(streams),2) as avg\_streams,

released\_month

from Spotify\_plst

where released\_year=2023

group by released\_month

order by released\_month

Select

CASE WHEN [danceability\_%] <= 25 THEN '0-25%'

WHEN [danceability\_%] <= 50 THEN '26-50%'

WHEN [danceability\_%] <= 75 THEN '51-75%'

ELSE '76-100%' END AS danceability\_group,

avg(streams)as avg\_streams

FROM Spotify\_plst

group by CASE WHEN [danceability\_%] <= 25 THEN '0-25%'

WHEN [danceability\_%] <= 50 THEN '26-50%'

WHEN [danceability\_%] <= 75 THEN '51-75%'

ELSE '76-100%' END

order by avg\_streams desc

/\*

select danceability\_group,

avg(streams) as avg\_streams from

(Select CASE WHEN [danceability\_%] <= 25 THEN '0-25%'

WHEN [danceability\_%] <= 50 THEN '26-50%'

WHEN [danceability\_%] <= 75 THEN '51-75%'

ELSE '76-100%' END AS danceability\_group,

streams

FROM Spotify\_plst ) group by 1 \*/

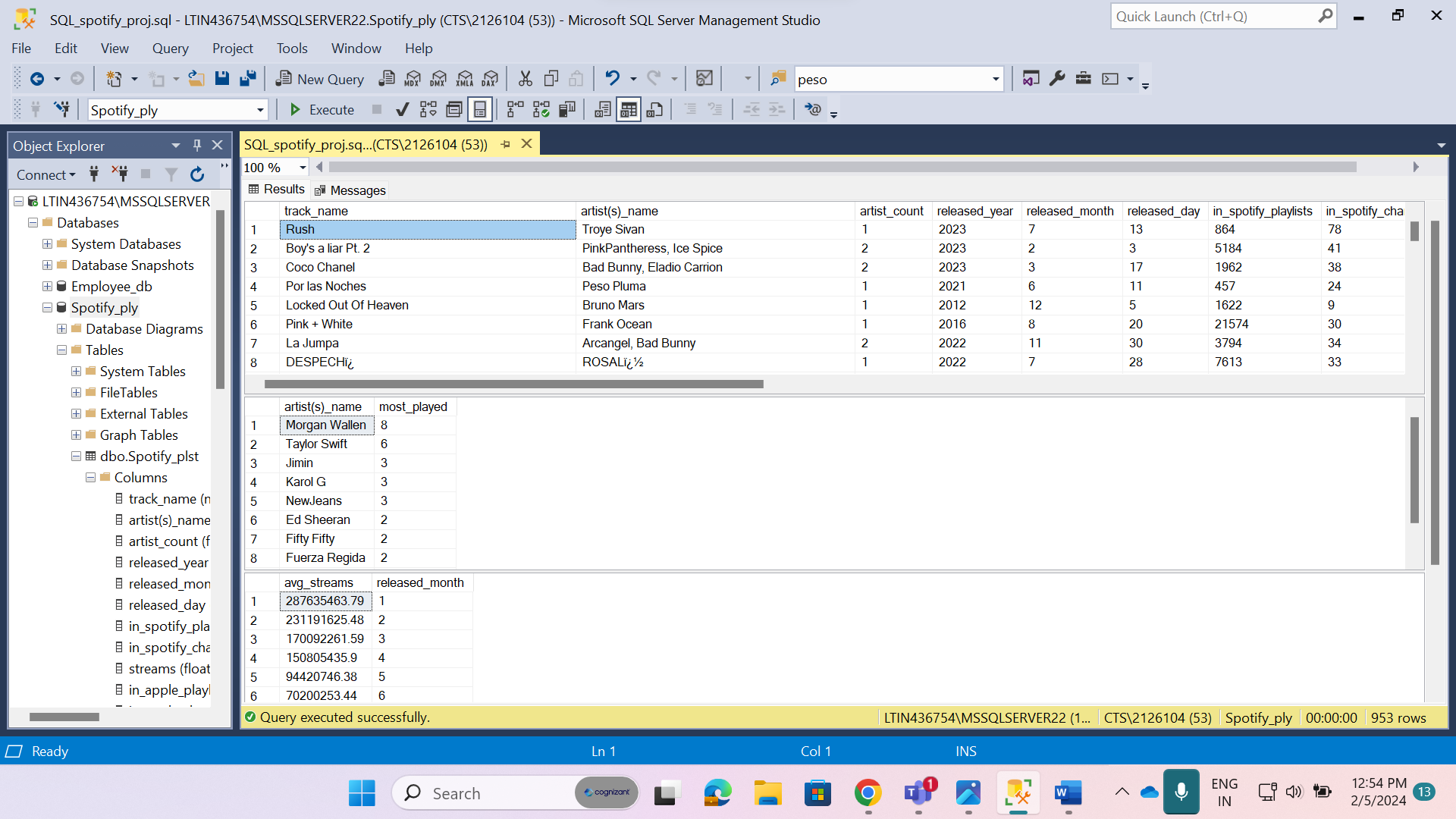
; with percent2023release as

(select count (\*) as total\_songs, sum(case when released\_year=2023 then 1 ELSE 0 END) as songs2023

from Spotify\_plst)

Select ((songs2023\*100)/total\_songs) as percent23

from percent2023release



* First result set depicts that whole data present in Spotify\_plst table.
* Second result depicts that the artist who has the most songs in the top streamed list in the year 2023

At the top – Morgan Wallen with 8 songs

* Third result depicts that average number of streams that has released in each month of year 2023

A screenshot of a computer

Description automatically generated

* Fourth result depicts that the (26-50%) danceability\_% group has more popular streams.
* Fifth result depicts that Top Spotify songs released in 2023 was 18% of songs released every year.