SPOTIFY SONGS DATA ANALYST PROJECT SQL

--Basic Level

1. Find the Total Streams by Genre

```
select genre,sum(stream)as total_streams
from spotifysongs
group by genre
order by total_streams;
```

	genre	total_streams
▶	Reggae	99854744176
	R&B	154259578872
	Jazz	248852257152
	Country	253438259344
	Classical	255875872220
	Folk	258405831704
	Hip-Hop	1235050550092
	Pop	1251263700542
	Electronic	1262182391256

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

2. Get the Top 10 Most Streamed Songs

```
select distinct song_title,artist,stream
from spotifysongs
order by stream desc
limit 10;
```

song_title	artist	stream
Development once.	Catherine Anderson	99999128
Lot religious.	Selena Perez	99998468
Certain why.	Thomas Roach	99996451
Next face.	Jason Strickland	99993859
Which hand relationship.	Monica Stevens	99992156
Speech.	Donna Campbell	99991525
Purpose cover.	Steven Cruz	99989128
Provide indeed.	Adam Marshall	99988733
Increase imagine.	Maria Davis	99988638
Establish high enough seem.	Donald Bennett	99986937

3. Count of Explicit vs. Non-Explicit Songs

select explicit_content, count(*) as Song_count

from spotifysongs

group by explicit_content;

explicit_content	Song_count
Yes	50166
No	49834

4. Find the Average Song Duration by Language

select language, avg(duration) as Avg_Duration_seconds

from spotifysongs

group by language;

5. Number of Songs Released Each Year

select release_date as Year, count(*) as Songs_Released
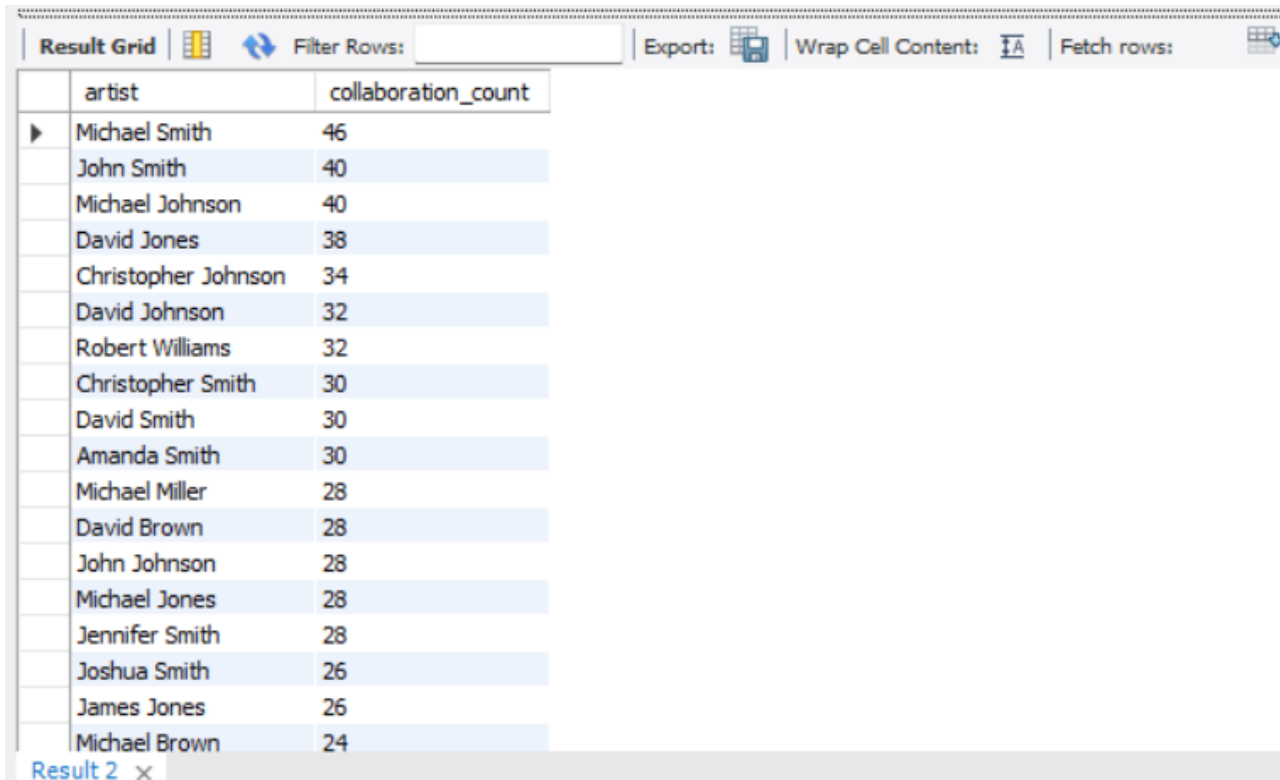from spotifysongs
group by Year
order by Year;

-- <u>Intermediate Level</u>

6. Artists with the Most Collaborations

select artist, count(*) as collaboration_count

from spotifysongs

group by artist

order by collaboration_count desc;

artist	collaboration_count
Michael Smith	46
John Smith	40
Michael Johnson	40
David Jones	38
Christopher Johnson	34
David Johnson	32
Robert Williams	32
Christopher Smith	30
David Smith	30
Amanda Smith	30
Michael Miller	28
David Brown	28
John Johnson	28
Michael Jones	28
Jennifer Smith	28
Joshua Smith	26
James Jones	26
Michael Brown	24

Result 2 ×

7. Top 5 Labels by Total Streams

select label, sum(stream) as total_Streams

from spotifysongs

group by label

order by total_Streams desc

limit 5;

label	total_Streams
Def Jam	848080773506
Sony Music	846547250904
Indie	842168687310
Atlantic Records	832722691790
Warner Music	830604340962

8. Most Prolific Composers (Top 10)

select composer, count(distinct song_id) as Songs_Composed

from spotifysongs

group by composer

order by Songs_Composed desc

limit 10;

composer	Songs_Composed
Michael Smith	21
Michael Jones	21
Michael Williams	20
Robert Smith	19
John Smith	18
James Jones	17
Michael Davis	16
Jennifer Smith	15
Jennifer Johnson	14
Brian Smith	14

9. Genre Popularity Trend Over Time Years

select Release_date as Year, genre, sum(stream) as Streams

from spotifysongs
group by Year, genre
order by Year, Streams desc;

Year	genre	Streams
01-01-1995	Hip-Hop	86573568
01-01-1995	Electronic	49147118
01-01-1996	Electronic	112277718
01-01-1996	Pop	20499400
01-01-1997	Electronic	282210714
01-01-1997	Pop	92164590
01-01-1997	Hip-Hop	88349696
01-01-1998	Pop	192662548
01-01-1998	Jazz	176536348
01-01-1998	Electronic	139408186
01-01-1998	Hip-Hop	97635628
01-01-1998	R&B	69818662
01-01-1999	Hip-Hop	219393034
01-01-1999	R&B	96416728
01-01-1999	Jazz	90604508
01-01-1999	Country	25822846
01-01-2000	Pop	148702234
01-01-2000	Classical	92820370

Result 5 ✕

10. Find Artists with Songs in Multiple Languages

select artist, count(distinct language) as Languages_Count
from spotifysongs
group by artist
having count(distinct language) > 1
order by Languages_Count desc;

artist	Languages_Count
Amanda Smith	7
Christopher Smith	7
David Smith	6
James Smith	6
John Smith	6
Michael Brown	6
Robert Williams	6
Brian Adams	5
Cheryl Johnson	5
David Miller	5
Deborah Smith	5
James Jones	5
James Young	5
Jennifer Smith	5
John Williams	5
Joseph Brown	5
Kevin Smith	5
Mark Smith	5

Result 6 ×

11. Songs with Longest Duration per Genre

select distinct genre, song_title, artist, duration

from (

   select genre, song_title, artist, duration,

      rank() over (partition by genre order by duration desc) as rank_

   FROM spotifysongs

) as ranked_songs

where rank_ = 1;

genre	song_title	artist	duration
Classical	Teach choice.	Gary Newton	433
Country	Present follow say.	Alejandra Scott	411
Electronic	Father family.	Heather Lewis	418
Folk	Report mother knowledge.	Paul Holland	417
Hip-Hop	Receive discuss police.	Maria Ruiz	428
Jazz	Defense enjoy conference.	Brenda Simon	429
Pop	Fact product until.	David Walker	429
R&B	Possible agency.	Linda Henry	399
Reggae	Business voice age.	Sherry Simmons	405

-- <u>Advanced Level</u>

12. Year on year Growth Rate in Streams by Genre

with Yearly_Streams as (

   select distinct Genre, Release_date as Year, sum(Stream) as Streams

   from Spotifysongs

  group by Genre, Year

)

select distinct Genre, Year, Streams,

    (Streams - lag(Streams) over (partition by Genre order by Year)) /

lag(Streams)

    over (partition by Genre order by Year) * 100 as Growth_Percentage

from Yearly_Streams

order by Genre, Year;

Genre	Year	Streams	Growth_Percenta
Classical	01-01-2000	92820370	NULL
Classical	01-01-2002	59832016	-35.5400
Classical	01-01-2005	189873798	217.3448
Classical	01-01-2008	181219686	-4.5578
Classical	01-01-2010	37631734	-79.2342
Classical	01-01-2015	93884048	149.4811
Classical	01-01-2023	134296056	43.0446
Classical	01-02-1999	60467460	-54.9745
Classical	01-02-2001	174971822	189.3653
Classical	01-02-2002	151330720	-13.5114
Classical	01-02-2010	65242664	-56.8874
Classical	01-02-2020	153369644	135.0757
Classical	01-02-2022	225411430	46.9726
Classical	01-03-1999	151750914	-32.6783
Classical	01-03-2001	172297270	13.5395
Classical	01-03-2014	195009774	13.1822
Classical	01-03-2021	41135336	-78.9060
Classical	01-04-2001	123989370	201.4182

Result 8 ✕

13. Identify "Hit" Songs (Top 1% by Streams)

with Percentile AS (

   select distinct Song_title, Artist, Stream,

      ntile(100) over (order by Stream desc) as Percentile_Rank

   from Spotifysongs

)

select Song_title, Artist, Stream

from Percentile

where Percentile_Rank = 1;

	Song_title	Artist	Stream
▶	Development once.	Catherine Anderson	99999128
	Lot religious.	Selena Perez	99998468
	Certain why.	Thomas Roach	99996451
	Next face.	Jason Strickland	99993859
	Which hand relationship.	Monica Stevens	99992156
	Speech.	Donna Campbell	99991525
	Purpose cover.	Steven Cruz	99989128
	Provide indeed.	Adam Marshall	99988733
	Increase imagine.	Maria Davis	99988638
	Establish high enough seem.	Donald Bennett	99986937
	Natural write side.	Timothy Velazquez	99984245
	Participant since hot for.	Brandy Smith MD	99984048
	Even find ask condition.	Walter Cohen	99983249
	Less wife.	Timothy Fisher	99981514
	Detail eat person.	David Wolf	99981160
	Rise.	Garrett Collins	99978195
	Report world.	Kyle Carlson	99973809
	Drug wonder market.	Brian Conner	99973578

Result 9 ✕

14. Producer-Artist Pairs with Highest Success Rate

select Producer, Artist, avg(Stream) as Avg_Streams, count(*) as Songs_Produced

from Spotifysongs

group by Producer, Artist

having count(*) > 3

order by Avg_Streams desc ;

	Producer	Artist	Avg_Streams	Songs_Produced
▶	Angela Martin	Sarah Stewart	52427098.5000	4

15. Label Market Share by Genre

with Genre_Label_Streams as (

   select Genre, label, sum(Stream) as Total_Streams

   from Spotifysongs

  group by Genre, label

),

Genre_Totals as (

   select Genre, sum(Total_Streams) as Genre_Total

   from Genre_Label_Streams

   group by Genre

)

select gls.Genre, gls.label,

    (gls.Total_Streams / gt.Genre_Total) * 100 as Market_Share_Percentage

from Genre_Label_Streams gls

join Genre_Totals gt on gls.Genre = gt.Genre

order by gls.Genre, Market_Share_Percentage desc;

	Genre	label	Market_Share_Percentage
▶	Classical	Sony Music	18.2558
	Classical	Def Jam	17.8448
	Classical	Warner Music	17.5499
	Classical	Universal Music	16.0991
	Classical	Indie	15.3653
	Classical	Atlantic Records	14.8850
	Country	Atlantic Records	17.5073
	Country	Indie	16.8289
	Country	Sony Music	16.7228
	Country	Warner Music	16.5435
	Country	Def Jam	16.3229
	Country	Universal Music	16.0745
	Electronic	Def Jam	17.0627
	Electronic	Warner Music	16.9276
	Electronic	Indie	16.7542
	Electronic	Sony Music	16.4790
	Electronic	Universal Music	16.4064
	Electronic	Atlantic Records	16.3700