**[SQL Project – Music Store Data Analysis](https://github.com/Prathamesh36/SQL_MUSIC_STORE_ANALYSIS)**

Data Analysis Project for Music Store to help how they can optimize their business opportunities and to help answering business related questions.

[**Music Store Datasets**](https://github.com/Prathamesh36/SQL_MUSIC_STORE_ANALYSIS/tree/main/Music%20Store%20Data)

[album.csv](https://github.com/Prathamesh36/SQL_MUSIC_STORE_ANALYSIS/blob/main/Music%20Store%20Data/album.csv) [album2.csv](https://github.com/Prathamesh36/SQL_MUSIC_STORE_ANALYSIS/blob/main/Music%20Store%20Data/album2.csv)

[Artist.csv](https://github.com/Prathamesh36/SQL_MUSIC_STORE_ANALYSIS/blob/main/Music%20Store%20Data/artist.csv) [customer.csv](https://github.com/Prathamesh36/SQL_MUSIC_STORE_ANALYSIS/blob/main/Music%20Store%20Data/customer.csv)

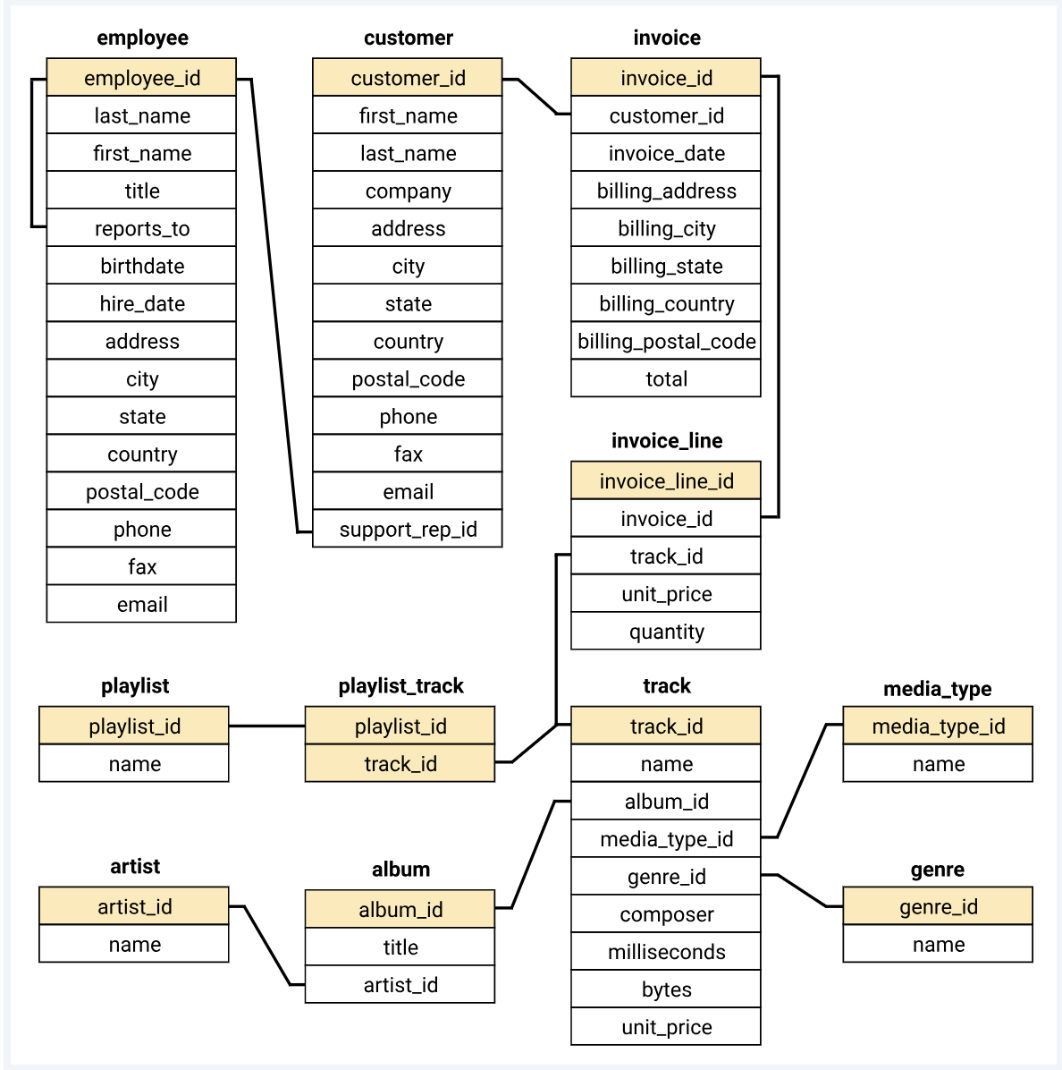
[Employee.csv](https://github.com/Prathamesh36/SQL_MUSIC_STORE_ANALYSIS/blob/main/Music%20Store%20Data/employee.csv) [genre.csv](https://github.com/Prathamesh36/SQL_MUSIC_STORE_ANALYSIS/blob/main/Music%20Store%20Data/genre.csv)

[Invoice.csv](https://github.com/Prathamesh36/SQL_MUSIC_STORE_ANALYSIS/blob/main/Music%20Store%20Data/invoice.csv) [invoice\_line.csv](https://github.com/Prathamesh36/SQL_MUSIC_STORE_ANALYSIS/blob/main/Music%20Store%20Data/invoice_line.csv)

[media\_type.csv](https://github.com/Prathamesh36/SQL_MUSIC_STORE_ANALYSIS/blob/main/Music%20Store%20Data/media_type.csv) [playlist.csv](https://github.com/Prathamesh36/SQL_MUSIC_STORE_ANALYSIS/blob/main/Music%20Store%20Data/playlist.csv)

[playlist\_track.csv](https://github.com/Prathamesh36/SQL_MUSIC_STORE_ANALYSIS/blob/main/Music%20Store%20Data/playlist_track.csv) [track.csv](https://github.com/Prathamesh36/SQL_MUSIC_STORE_ANALYSIS/blob/main/Music%20Store%20Data/track.csv)

**Schema Diagram**

****

**Database and Tools**

* PostgreSQL
* pgAdmin4

**Questions for the project**

[1. Who is the senior most employee based on job title?](#Q1)

[2. Which countries have the most Invoices?](#Q2)

[3. What are top 3 values of total invoice?](#Q3)

[4. Which city has the best customers? We would like to throw a promotional Music Festival in the city we made the most money. Write a query that returns one city that has the highest sum of invoice totals. Return both the city name & sum of all invoice totals](#Q4)

[5. Who is the best customer? The customer who has spent the most money will be declared the best customer. Write a query that returns the person who has spent the most money](#Q5)

[6. Write query to return the email, first name, last name, & Genre of all Rock Music listeners. Return your list ordered alphabetically by email starting with A](#Q6)

[7. Let's invite the artists who have written the most rock music in our dataset. Write a query that returns the Artist name and total track count of the top 10 rock bands](#Q7)

[8. Return all the track names that have a song length longer than the average song length. Return the Name and Milliseconds for each track. Order by the song length with the longest songs listed first](#Q8)

[9. Find how much amount spent by each customer on artists? Write a query to return customer name, artist name and total spent](#Q9)

[10. We want to find out the most popular music Genre for each country. We determine the most popular genre as the genre with the highest amount of purchases. Write a query that returns each country along with the top Genre. For countries where the maximum number of purchases is shared return all Genres](#Q10)

[11. Write a query that determines the customer that has spent the most on music for each country. Write a query that returns the country along with the top customer and how much they spent. For countries where the top amount spent is shared, provide all customers who spent this amount](#Q11)

**Questions and Answers (Queries)**

***1. Who is the senior most employee based on job title?***

* select first\_name, last\_name, levels from employee order by levels desc limit 1;



***2. Which countries have the most Invoices?***

* select billing\_country,count(\*) as Total\_count from invoice

group by billing\_country

order by Total\_count desc;

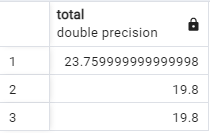


***3. What are top 3 values of total invoice?***

* select total from invoice

order by total desc

limit 3;



***4. Which city has the best customers? We would like to throw a promotional Music Festival in the city we made the most money. Write a query that returns one city that has the highest sum of invoice totals. Return both the city name & sum of all invoice totals***

* select billing\_city as city, sum(total) as total\_sum from invoice

group by city

order by total\_sum desc

limit 1;



***5. Who is the best customer? The customer who has spent the most money will be declared the best customer. Write a query that returns the person who has spent the most money***

* select c.customer\_id, c.first\_name, c.last\_name, sum(i.total)as money\_spent

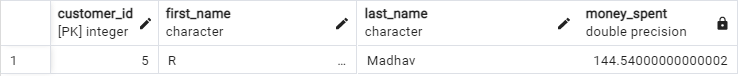
from customer as c

join invoice as I on c.customer\_id=i.customer\_id

group by c.customer\_id

order by money\_spent desc

limit 1;



***6. Write query to return the email, first name, last name, & Genre of all Rock Music listeners. Return your list ordered alphabetically by email starting with A***

* select distinct email,first\_name, last\_name

from customer

join invoice on customer.customer\_id=invoice.customer\_id

join invoice\_line on invoice.invoice\_id=invoice\_line.invoice\_id

where track\_id in(

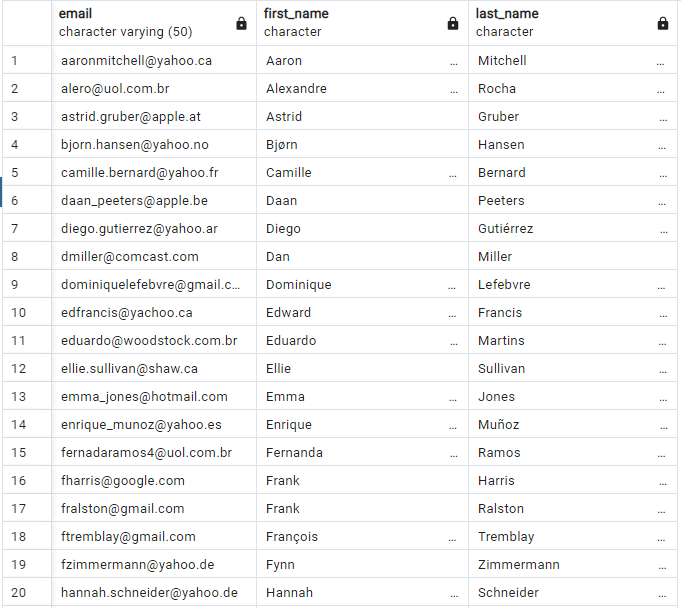
select track\_id from track

join genre on track.genre\_id=genre.genre\_id

where genre.name = 'Rock'

)

order by email ;

\*top 20 records only

***7. Let's invite the artists who have written the most rock music in our dataset. Write a query that returns the Artist name and total track count of the top 10 rock bands***

* select artist.artist\_id,artist.name, count(artist.artist\_id)as total\_count from artist

join album on artist.artist\_id=album.artist\_id

join track on album.album\_id=track.album\_id

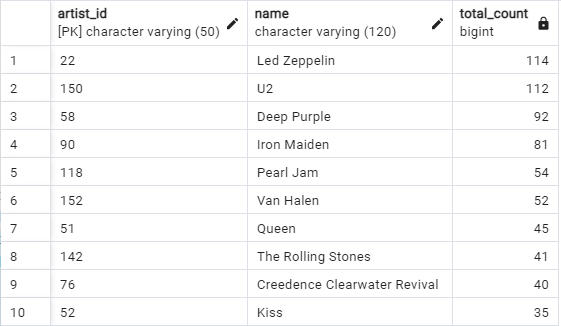
join genre on track.genre\_id=genre.genre\_id

where genre.name ='Rock'

group by artist.artist\_id

order by total\_count desc

limit 10;



***8. Return all the track names that have a song length longer than the average song length. Return the Name and Milliseconds for each track. Order by the song length with the longest songs listed first***

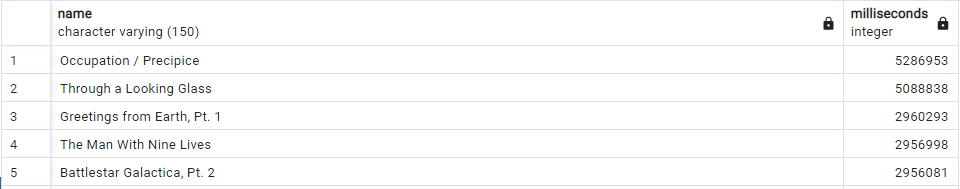
* select name, milliseconds from track

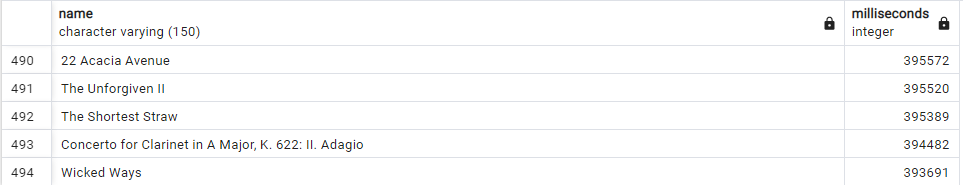
where milliseconds >(

select avg(milliseconds) from track

)

order by milliseconds desc;

…



\*top 5 and bottom 5 records

***9. Find how much amount spent by each customer on artists? Write a query to return customer name, artist name and total spent***

* with best\_selling\_artist AS (

SELECT artist.artist\_id AS artist\_id, artist.name AS artist\_name,

SUM(invoice\_line.unit\_price\*invoice\_line.quantity) AS total\_sales

FROM invoice\_line

JOIN track on track.track\_id = invoice\_line.track\_id

JOIN album ON album.album\_id = track.album\_id

JOIN artist ON artist.artist\_id = album.artist\_id

Group BY 1

ORDER BY 3 DESC

LIMIT 1

)

SELECT c.customer\_id,c.first\_name, c.last\_name, bsa.artist\_name,

SUM (il.unit\_price\*il.quantity) As amount\_spent

FROM invoice i

JOIN customer c ON c.customer\_id = i.customer\_id

JOIN invoice\_line il ON il.invoice\_id = i.invoice\_id

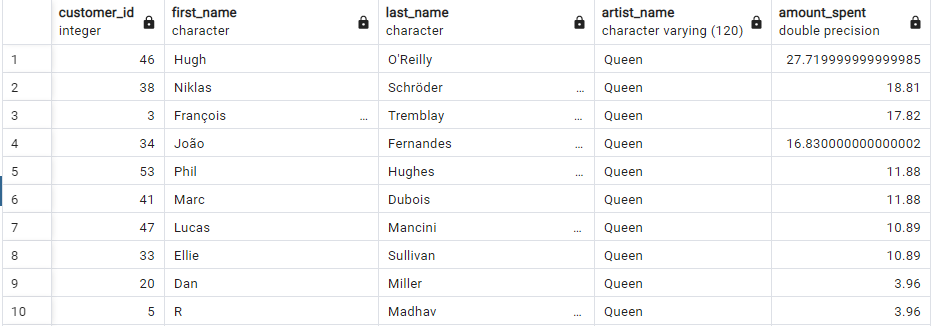
JOIN track t on t.track\_id =il.track\_id

JOIN album alb ON alb.album\_id = t.album\_id

JOIN best\_selling\_artist bsa ON bsa.artist\_id =alb.artist\_id

GROUP BY 1,2,3,4

ORDER BY 5 DESC;

\*top 10 records only

***10. We want to find out the most popular music Genre for each country. We determine the most popular genre as the genre with the highest amount of purchases. Write a query that returns each country along with the top Genre. For countries where the maximum number of purchases is shared return all Genres***

* WITH popular\_genre AS (

SELECT COUNT(invoice\_line.quantity) AS purchases, customer.country, genre.name,genre.genre\_id,

ROW\_NUMBER() OVER ( PARTITION BY customer.country ORDER BY COUNT (invoice\_line.quantity) DESC) AS RowNo

FROM invoice\_line

JOIN invoice ON invoice.invoice\_id = invoice\_line.invoice\_id

JOIN customer ON customer.customer\_id =invoice.customer\_id

JOIN track ON track.track\_id = invoice\_line.track\_id

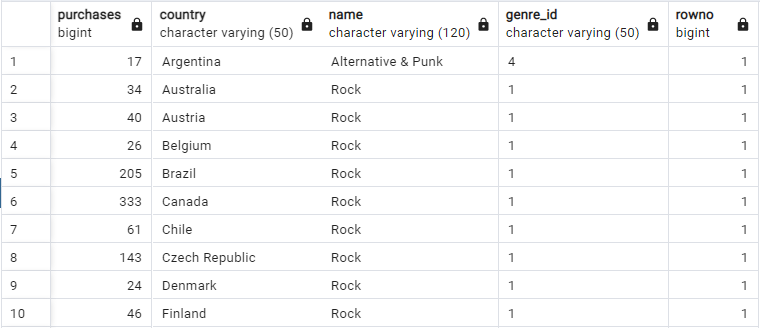
JOIN genre ON genre.genre\_id = track.genre\_id

GROUP by 2,3,4

ORDER BY 2 ASC , 1 DESC

)

SELECT \* from popular\_genre WHERE RowNo <= 1



\*top 10 records only

***11. Write a query that determines the customer that has spent the most on music for each country. Write a query that returns the country along with the top customer and how much they spent. For countries where the top amount spent is shared, provide all customers who spent this amount***

* with Customer\_with\_country AS (

SELECT customer.customer\_id,first\_name, last\_name,billing\_country,SUM(total) AS total\_spending,

ROW\_NUMBER() OVER (PARTITION BY billing\_country ORDER BY SUM(total) DESC) AS RowNo

FROM invoice

JOIN customer ON customer.customer\_id = invoice.customer\_id

GROUP BY 1,2,3,4

ORDER BY 4 ASC, 5 DESC

)

SELECT \* FROM Customer\_with\_Country WHERE RowNo <= 1;

