

SQL PROJECT

MUSIC STORE DATA ANALYSIS

**Created by -
Akash Singh**

OBJECTIVE



This analysis aims to increase the business growth of the music store by addressing some existing challenges.

We want to help the music store grow by looking at its data. By using SQL and answering some questions, we can figure out how to tackle the challenges and boost the business.

Let's dive in...



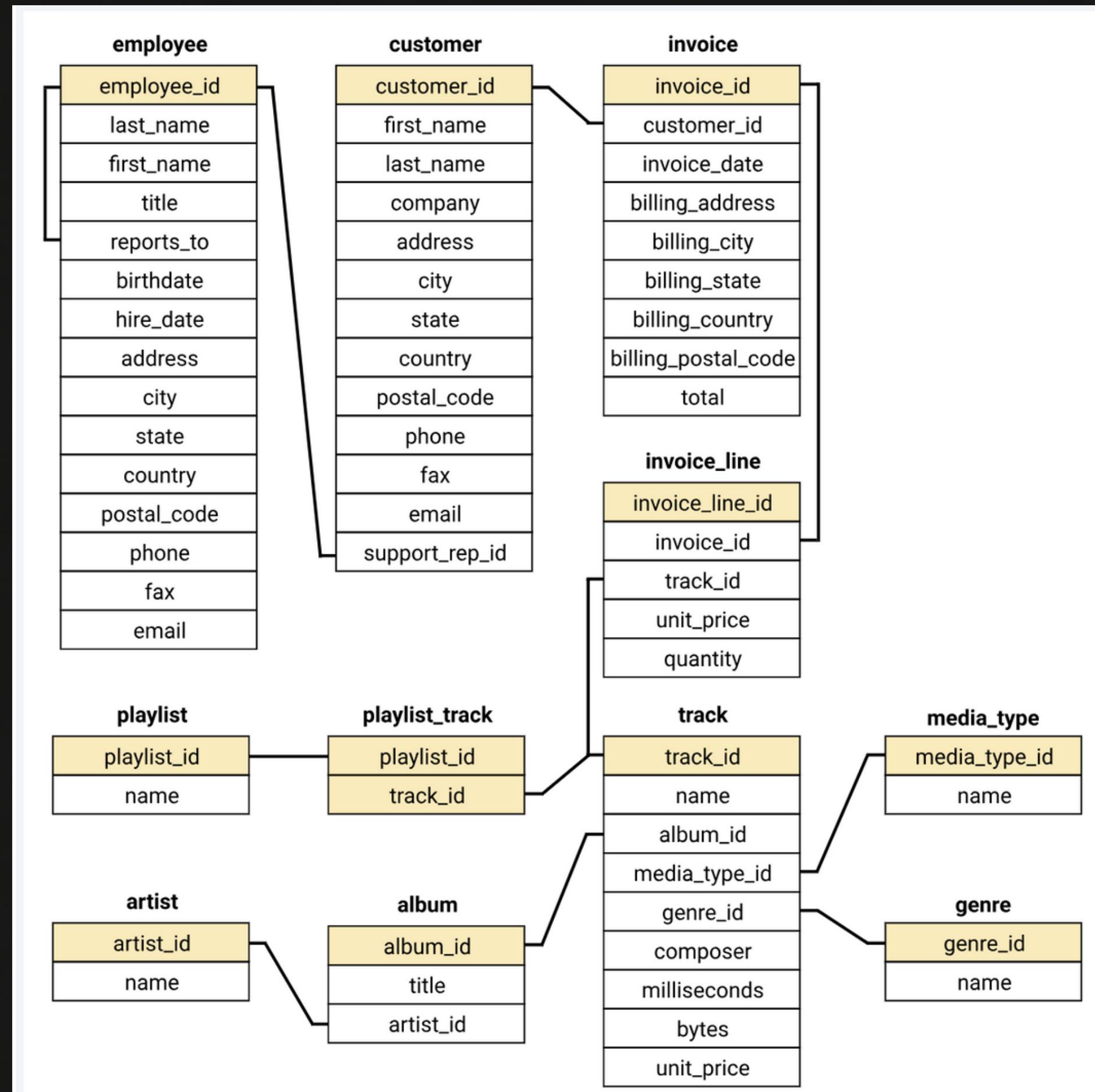
QUESTION DIVISION

**Questions are divided into 3 categories
based on their complexity-**

- 1. Easy**
- 2. Moderate**
- 3. Advance**



DATABASE SCHEMA



EASY

Q1: Who is the senior most employee based on job title?

↓ Ans

```
SELECT *
FROM employee
ORDER BY levels DESC
LIMIT 1
```

↓ Output

employee_id	last_name	first_name	title	reports_to	levels	
1	Adams	Andrew	General Manager	9	L6	1



EASY

Q2: Which countries have the most Invoices?

↓ Ans

```
SELECT COUNT(*) AS c, billing_country  
FROM invoice  
GROUP BY billing_country  
ORDER BY c DESC
```

Output

→

c	billing_country
131	USA
76	Canada
61	Brazil
50	France
41	Germany
30	Czech Republic
29	Portugal
28	United Kingdom
21	India



EASY

Q3: What are the top 3 values of the total invoice?

↓ Ans

```
SELECT total  
FROM invoice  
ORDER BY total DESC  
limit 3
```

↓ Output

total
23.759999999999998
19.8
19.8



EASY

Q4: Which city has the best customers? We would like to throw a promotional Music Festival in the city where 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

↓ Ans

```
SELECT billing_city, SUM(total) AS InvoiceTotal  
FROM invoice  
GROUP BY billing_city  
ORDER BY InvoiceTotal DESC  
LIMIT 1;
```

↓ Output

billing_city	InvoiceTotal
Prague	273.2400000000007



EASY

Q5: 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.

↓ Ans

```
SELECT customer.customer_id, first_name, last_name, SUM(total) AS total_spending  
FROM customer  
JOIN invoice ON customer.customer_id = invoice.customer_id  
GROUP BY customer.customer_id  
ORDER BY total_spending DESC  
LIMIT 1;
```

↓ Output

customer_id	first_name	last_name	total_spending
5	František	Wichterlová	144.54000000000002



MODERATE

Q1: Write a query to return the email, first name, last name, and genre of all Rock Music listeners. Return your list ordered alphabetically by email starting with A.

↓ Ans

```
-- Method 1
SELECT DISTINCT first_name, last_name, email, genre.name
FROM customer
JOIN invoice
ON customer.customer_id = invoice.customer_id
JOIN invoice_line
ON invoice.invoice_id = invoice_line.invoice_id
JOIN track
ON invoice_line.track_id = track.track_id
JOIN genre
ON track.genre_id = genre.genre_id
WHERE genre.name like "Rock"
```

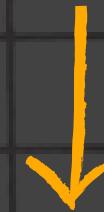
Output →

first_name	last_name	email	name
Martha	Silk	marthasilk@gmail.com	Rock
StanisÅ	StanisÅ,aw	stanisÅ,aw.wÅ³jcik@wp.pl	Rock
Wyatt	Girard	wyatt.girard@yahoo.fr	Rock
LuÅ-s	GonÅsalves	luisg@embraer.com.br	Rock
Dan	Miller	dmiller@comcast.com	Rock
Isabelle	Mercier	isabelle_mercier@apple.fr	Rock
Kara	Nielsen	kara.nielsen@jubii.dk	Rock
Phil	Hughes	phil.hughes@gmail.com	Rock
Camille	Bernard	camille.bernard@yahoo.fr	Rock
Mark	Taylor	mark.taylor@yahoo.au	Rock



MODERATE

Q1: Write a query to return the email, first name, last name, and genre of all Rock Music listeners. Return your list ordered alphabetically by email starting with A.

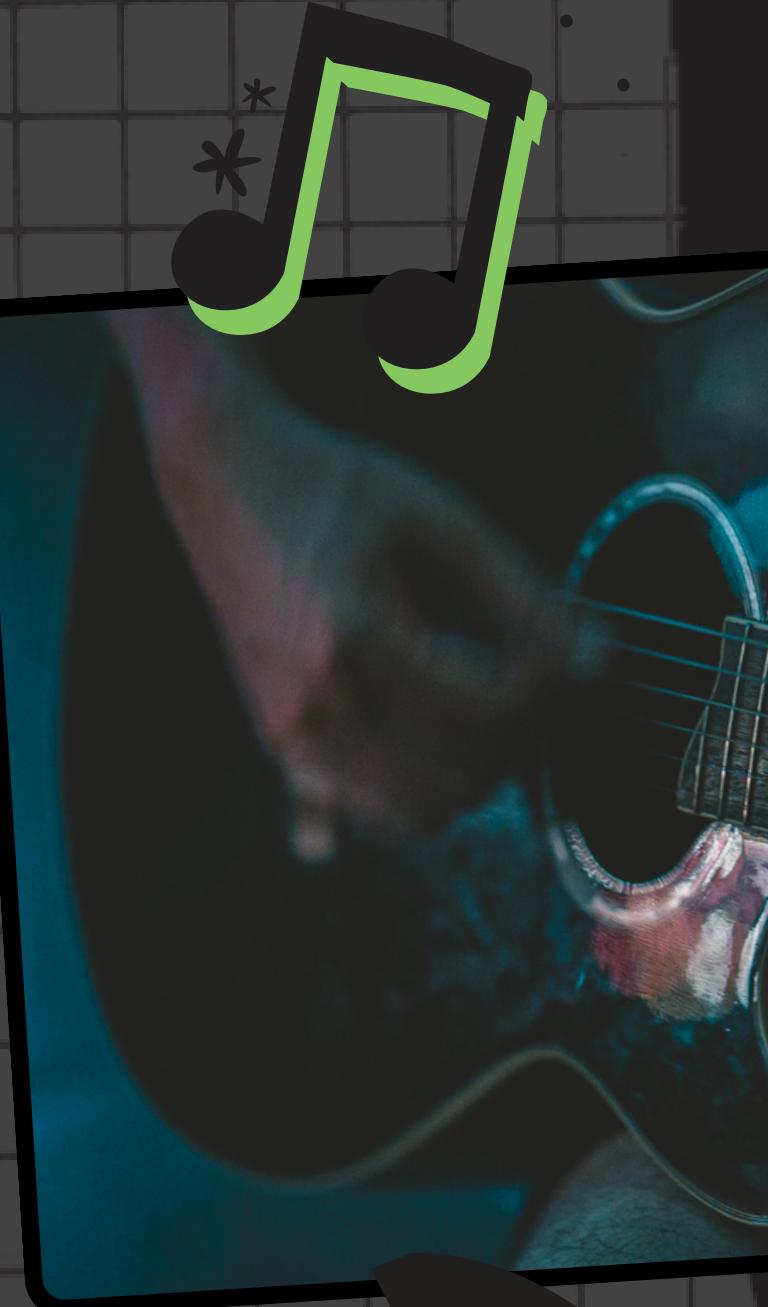


Ans

```
-- Method 2
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 LIKE 'Rock'
)
ORDER BY email;
```

Output →

email	first_name	last_name
aaronmitchell@yahoo.ca	Aaron	Mitchell
alero@uol.com.br	Alexandre	Rocha
astrid.gruber@apple.at	Astrid	Gruber
bjorn.hansen@yahoo.no	Bjørn	Hansen
camille.bernard@yahoo.fr	Camille	Bernard
daan_peeters@apple.be	Daan	Peeters
diego.gutierrez@yahoo.ar	Diego	Gutiérrez
dmiller@comcast.com	Dan	Miller
dominiquelefebvre@gmail.com	Dominique	Lefebvre
edfrancis@yahoo.ca	Edward	Francis



MODERATE

Q2: 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.



Ans

```
SELECT artist.artist_id, artist.name, COUNT(artist.artist_id) AS number_of_songs  
FROM track  
JOIN album ON album.album_id = track.album_id  
JOIN artist ON artist.artist_id = album.artist_id  
JOIN genre ON genre.genre_id = track.genre_id  
WHERE genre.name LIKE 'Rock'  
GROUP BY artist.artist_id, artist.name  
ORDER BY number_of_songs DESC  
LIMIT 10;
```

Output →

artist_id	name	number_of_songs
1	AC/DC	18
3	Aerosmith	15
8	Audioslave	14
22	Led Zeppelin	14
4	Alanis Morissette	13
5	Alice In Chains	12
23	Frank Zappa & Captain Beefheart	9
2	Accept	4



MODERATE

Q3: 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.

Ans

```
SELECT name,milliseconds
FROM track
WHERE milliseconds > (
    SELECT AVG(milliseconds) AS avg_track_length
    FROM track )
ORDER BY milliseconds DESC;
```

Output →

name	milliseconds
How Many More Times	711836
Advance Romance	677694
Sleeping Village	644571
You Shook Me(2)	619467
Talkin' 'Bout Women Obviously	589531
Stratus	582086
No More Tears	555075
The Alchemist	509413
Wheels Of Confusion / The Straightener	494524
Book Of Thel	494393
You Oughta Know (Alternate)	491885



ADVANCE

Q1: Find how much amount spent by each customer on artists? Write a query to return customer name, artist name and total spent



Ans

```
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 1,2
    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;
```

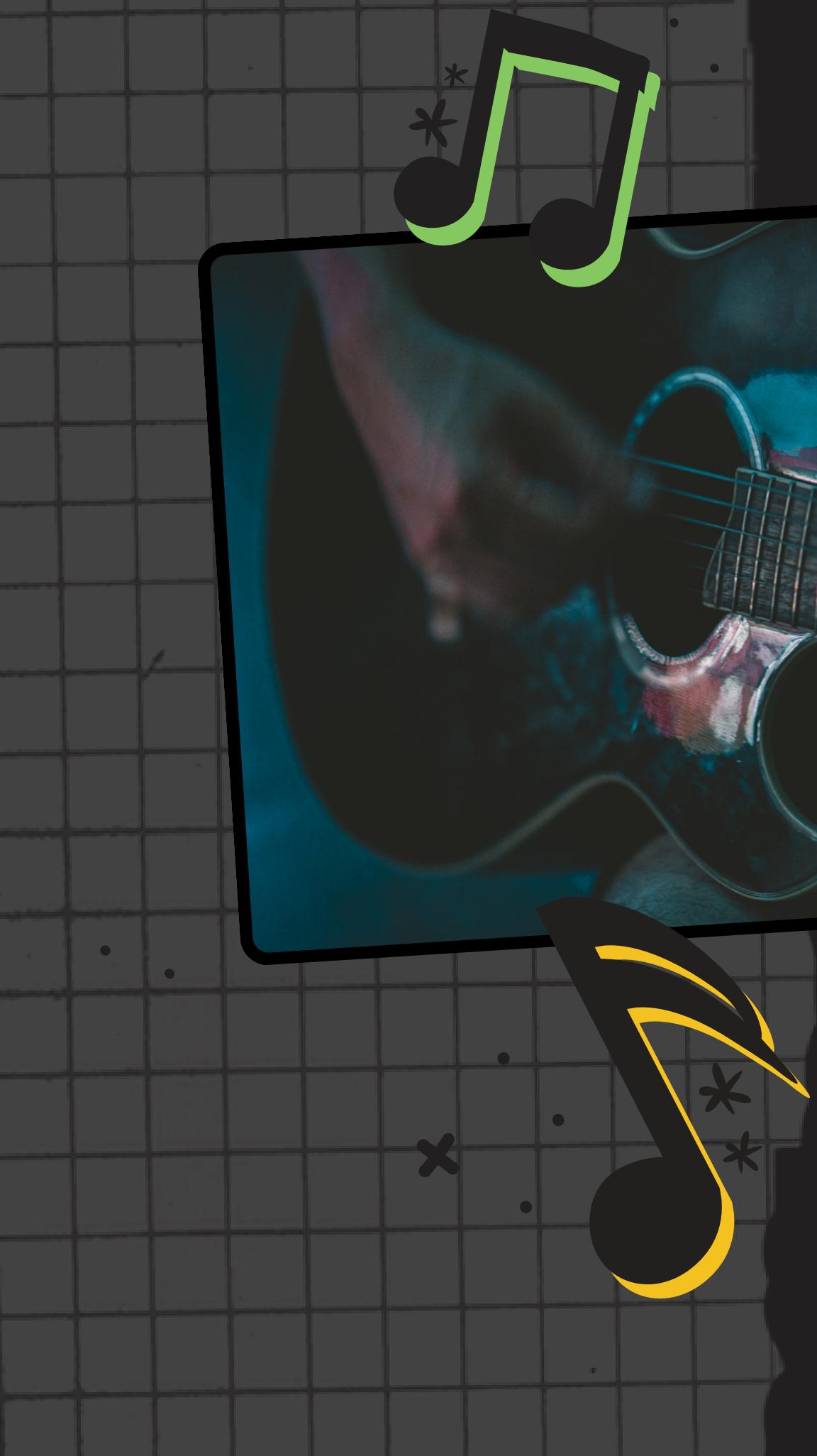


ADVANCE

Output



customer_id	first_name	last_name	artist_name	amount_spent
54	Steve	Murray	AC/DC	17.82
53	Phil	Hughes	AC/DC	10.89
21	Kathy	Chase	AC/DC	10.89
49	StanisÅaw	WÃ³jciech	AC/DC	9.9
1	LuÃ-s	GonÃ§alves	AC/DC	7.920000000000001
24	Frank	Ralston	AC/DC	7.920000000000001
31	Martha	Silk	AC/DC	3.96
44	Terhi	HÃ¤mÃ¤linen	AC/DC	2.9699999999999998
16	Frank	Harris	AC/DC	2.9699999999999998
42	Wyatt	Girard	AC/DC	2.9699999999999998
38	Niklas	SchrÃ¶der	AC/DC	2.9699999999999998
35	Madalena	Sampaio	AC/DC	2.9699999999999998
6	Helena	HolÃ½	AC/DC	2.9699999999999998



ADVANCE

Q2: 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.

↓ Ans

```
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
```



ADVANCE

Output



purchases	country	name	genre_id	RowNo
1	Argentina	Rock	1	1
18	Australia	Rock	1	1
6	Austria	Rock	1	1
5	Belgium	Rock	1	1
26	Brazil	Rock	1	1
57	Canada	Rock	1	1
7	Chile	Rock	1	1
14	Czech Republic	Rock	1	1
6	Denmark	Rock	1	1
6	Finland	Rock	1	1
26	France	Rock	1	1
28	Germany	Rock	1	1
4	Hungary	Rock	1	1
13	India	Rock	1	1
7	Ireland	Rock	1	1



ADVANCE

Q3: 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

↓ Ans

```
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
```



ADVANCE

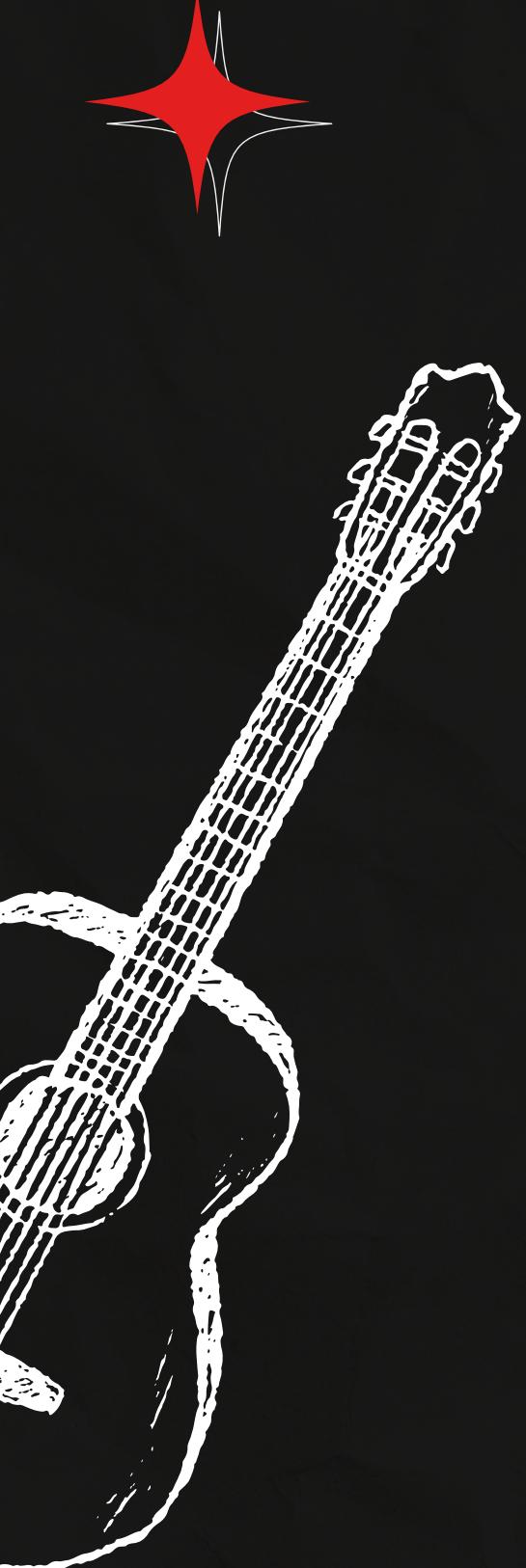
Output



customer_id	first_name	last_name	billing_country	total_spending	RowNo
56	Diego	Gutiérrez	Argentina	39.6	1
55	Mark	Taylor	Australia	81.18	1
7	Astrid	Gruber	Austria	69.3	1
8	Daan	Peeters	Belgium	60.38999999999999	1
1	Luís	Gonçalves	Brazil	108.89999999999998	1
3	François	Tremblay	Canada	99.99	1
57	Luis	Rojas	Chile	97.02000000000001	1
5	František	Wichterlová	Czech Republic	144.54000000000002	1
9	Kara	Nielsen	Denmark	37.61999999999999	1
44	Terhi	Häkämäki	Finland	79.2	1
42	Wyatt	Girard	France	99.99	1
37	Fynn	Zimmermann	Germany	94.05000000000001	1
45	Ladislav	Kovács	Hungary	78.21	1
58	Manoj	Pareek	India	111.86999999999999	1
46	Hinch	O'Reilly	Ireland	114.83999999999997	1



THANK YOU



watch
play
at
or
s-pl
f big
me y
nich

