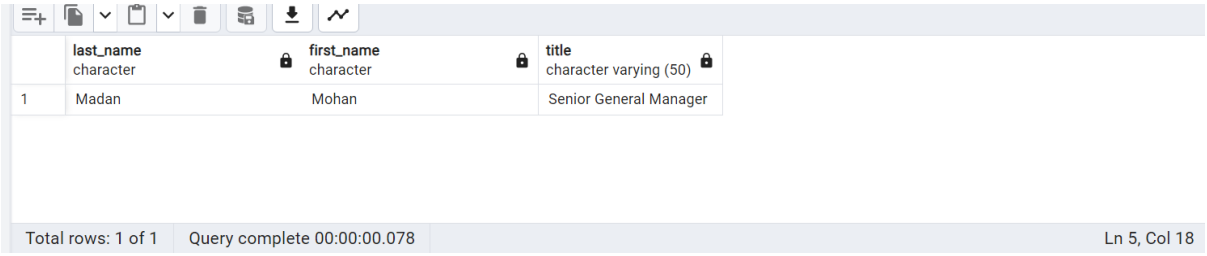
MUSIC STORE ANALYSIS IN SQL

SQL QUERY AND OUTPUT

/* Question Set 1 - Easy */

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

```
SELECT title, last_name, first_name
FROM employee
ORDER BY levels DESC
LIMIT 1
```

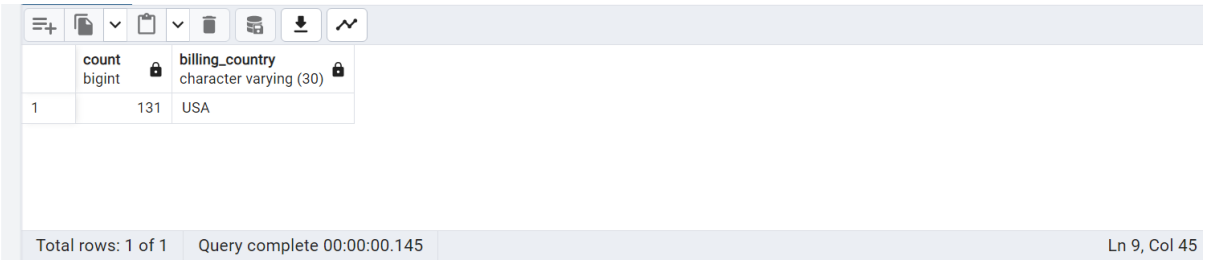


	last_name character	first_name character	title character varying (50)
1	Madan	Mohan	Senior General Manager

Total rows: 1 of 1 | Query complete 00:00:00.078 | Ln 5, Col 18

/* Q2: Which countries have the most Invoices? */

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



	count bigint	billing_country character varying (30)
1	131	USA

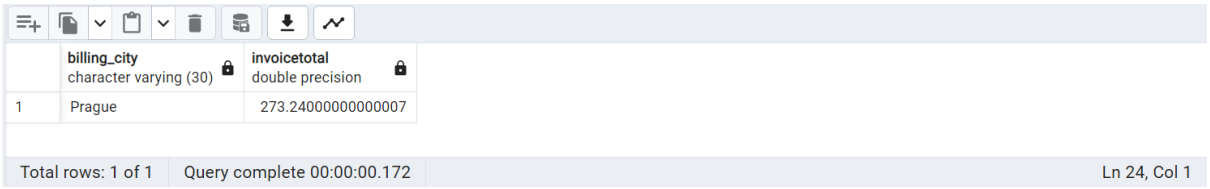
Total rows: 1 of 1 | Query complete 00:00:00.145 | Ln 9, Col 45

/* Q3: What are top 3 values of total invoice? */

SELECT total

FROM invoice

ORDER BY total DESC



	billing_city character varying (30)	invoicetotal double precision
1	Prague	273.24000000000007

Total rows: 1 of 1 | Query complete 00:00:00.172 | Ln 24, Col 1

/* Q4: 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 */

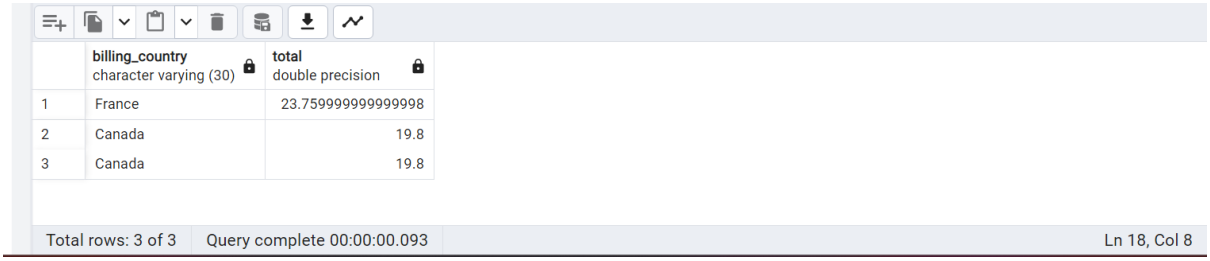
SELECT billing_city,SUM(total) AS InvoiceTotal

FROM invoice

GROUP BY billing_city

ORDER BY InvoiceTotal DESC

LIMIT 1;



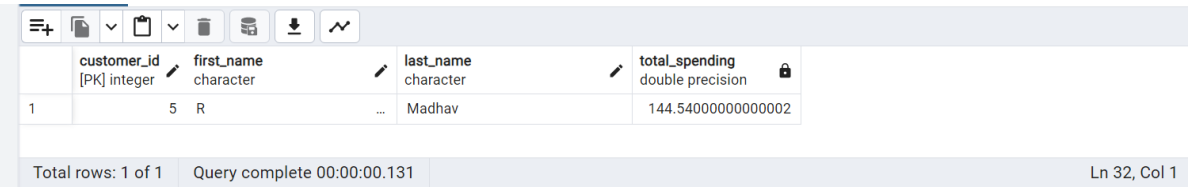
	billing_country character varying (30)	total double precision
1	France	23.759999999999998
2	Canada	19.8
3	Canada	19.8

Total rows: 3 of 3 | Query complete 00:00:00.093 | Ln 18, Col 8

/* 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.*/

```
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;
```



	customer_id [PK] integer	first_name character	last_name character	total_spending double precision
1	5	R ...	Madhav	144.54000000000002

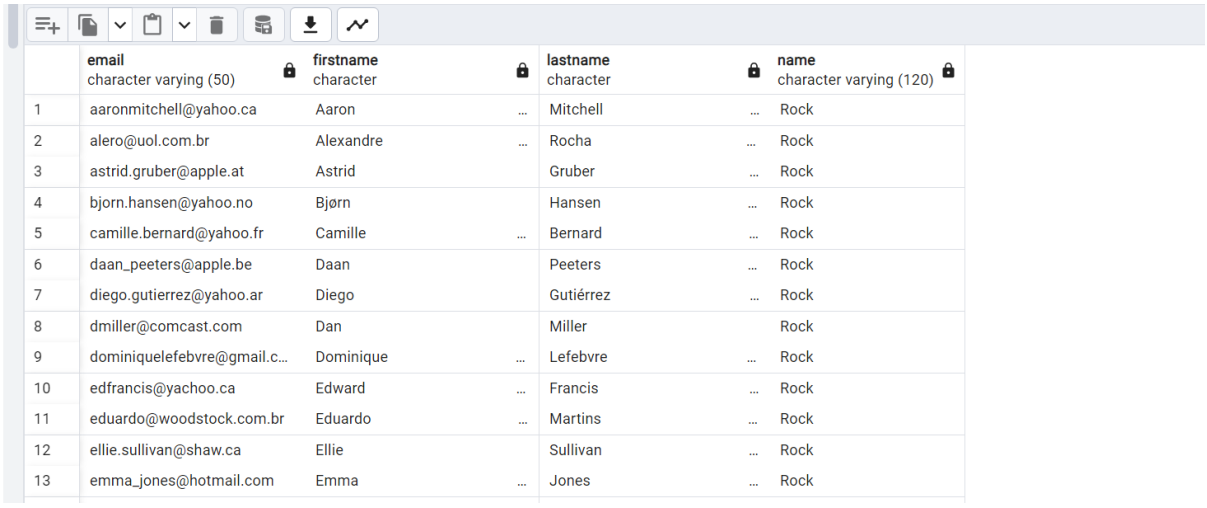
Total rows: 1 of 1 | Query complete 00:00:00.131 | Ln 32, Col 1

/* Question Set 2 - Moderate */

/* Q1: 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 AS Email,first_name AS FirstName, last_name AS LastName, genre.name AS Name
FROM customer
JOIN invoice ON invoice.customer_id = customer.customer_id
JOIN invoice_line ON invoice_line.invoice_id = invoice.invoice_id
JOIN track ON track.track_id = invoice_line.track_id
JOIN genre ON genre.genre_id = track.genre_id
WHERE genre.name LIKE 'Rock'
ORDER BY email;
```

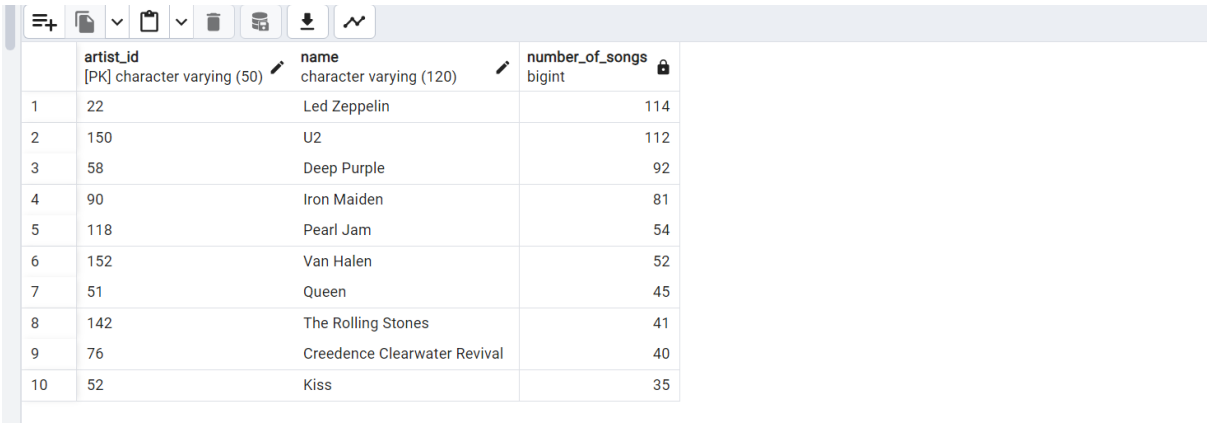


	email character varying (50)	firstname character	lastname character	name character varying (120)
1	aaronmitchell@yahoo.ca	Aaron	Mitchell	Rock
2	alero@uol.com.br	Alexandre	Rocha	Rock
3	astrid.gruber@apple.at	Astrid	Gruber	Rock
4	bjorn.hansen@yahoo.no	Bjørn	Hansen	Rock
5	camille.bernard@yahoo.fr	Camille	Bernard	Rock
6	daan_peeters@apple.be	Daan	Peeters	Rock
7	diego.gutierrez@yahoo.ar	Diego	Gutiérrez	Rock
8	dmiller@comcast.com	Dan	Miller	Rock
9	dominiquelefebvre@gmail.c...	Dominique	Lefebvre	Rock
10	edfrancis@yachoo.ca	Edward	Francis	Rock
11	eduardo@woodstock.com.br	Eduardo	Martins	Rock
12	ellie.sullivan@shaw.ca	Ellie	Sullivan	Rock
13	emma_jones@hotmail.com	Emma	Jones	Rock

/* 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.
***/**

```
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
ORDER BY number_of_songs DESC
LIMIT 10;
```



	artist_id [PK] character varying (50)	name character varying (120)	number_of_songs bigint
1	22	Led Zeppelin	114
2	150	U2	112
3	58	Deep Purple	92
4	90	Iron Maiden	81
5	118	Pearl Jam	54
6	152	Van Halen	52
7	51	Queen	45
8	142	The Rolling Stones	41
9	76	Creedence Clearwater Revival	40
10	52	Kiss	35

/* 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.*/

```
select name ,milliseconds  from track

where  milliseconds>(select avg(milliseconds)

                     from track)

order by milliseconds desc
```

	name character varying (150)	milliseconds integer
1	Occupation / Precipice	5286953
2	Through a Looking Glass	5088838
3	Greetings from Earth, Pt. 1	2960293
4	The Man With Nine Lives	2956998
5	Battlestar Galactica, Pt. 2	2956081
6	Battlestar Galactica, Pt. 1	2952702
7	Murder On the Rising Star	2935894
8	Battlestar Galactica, Pt. 3	2927802
9	Take the Celestra	2927677
10	Fire In Space	2926593
11	The Long Patrol	2925008
12	The Magnificent Warriors	2924716
13	The Living Legend, Pt. 1	2924507
14	The Gun On Ice Planet Zero, Pt. 2	2924341
15	The Hand of God	2924007
16	Experiment In Terra	2923548

Total rows: 494 of 494 | Query complete 00:00:00.657 | Ln 85, Col 27

/* Question Set 3 - Advance */

/* Q1: Find how much amount spent by each customer on artists?

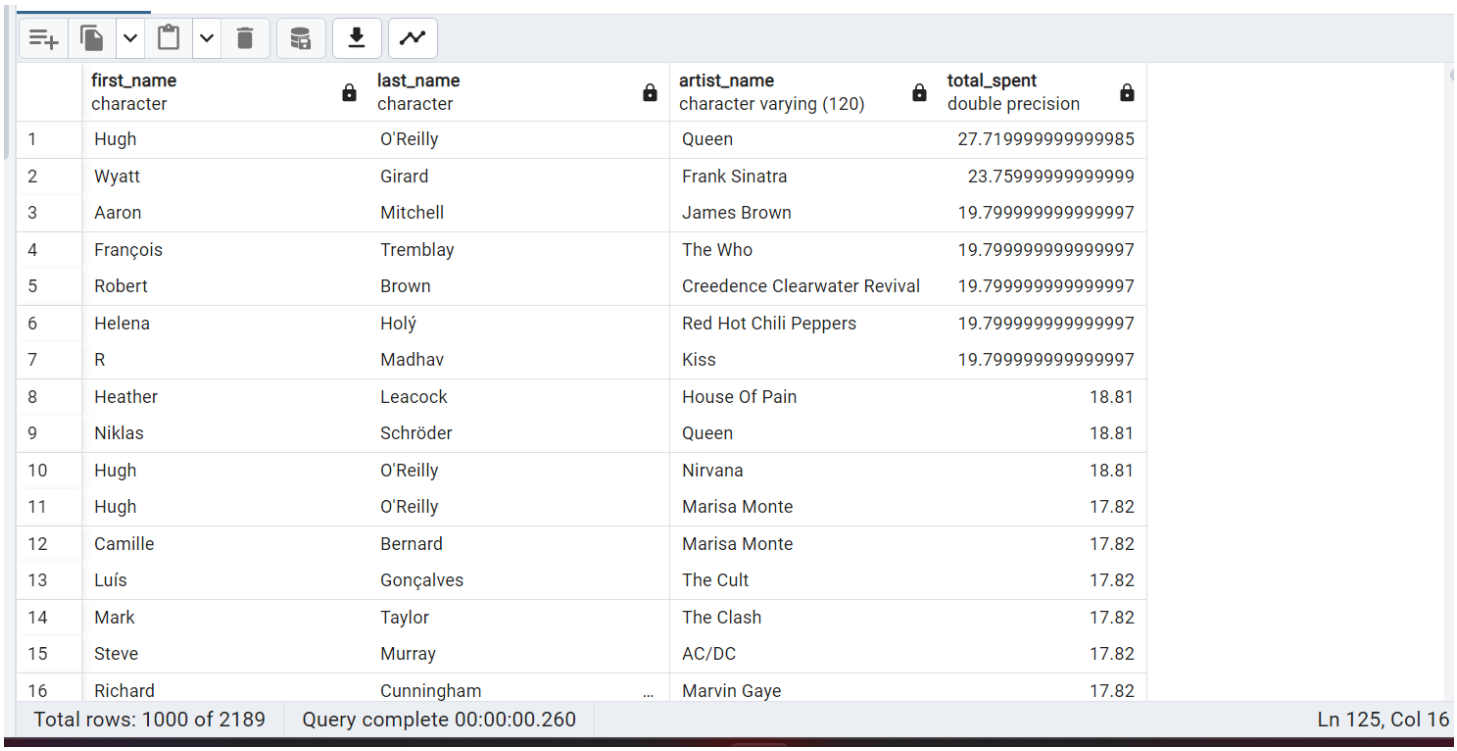
Write a query to return customer name, artist name and total spent */

/*Steps to solve:customer name is in the "customer" table,artist name is in the "artist" table

and I have calculated total_sales on each song from invoice table by multiplying unitprice and quantity.With the help of common table query(CTE) the idea is to create a temporary table containing coumns of artist_name and total_sales.Through this table it can be inferred how much amount is spent on each artist by all customer.Now to add customer name;customer table is joined with invoice table and then consecutively it is joined with invoice_line,track,album,artist and best_selling_artist(bsa) table to get artist name along with customer name */

```
with best_selling_artist as (
        select 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 album.artist_id=artist.artist_id
        group by artist.name
        order by total_sales desc)
select customer.first_name,customer.last_name,bsa.artist_name,
sum(invoice_line.unit_price*invoice_line.quantity)as total_spent
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 track.track_id=invoice_line.track_id
join album on album.album_id=track.album_id
join artist on album.artist_id=artist.artist_id
join best_selling_artist bsa on bsa.artist_name=artist.name
group by 1,2,3
order by 4 desc
```

O/P



	first_name character	last_name character	artist_name character varying (120)	total_spent double precision
1	Hugh	O'Reilly	Queen	27.719999999999985
2	Wyatt	Girard	Frank Sinatra	23.75999999999999
3	Aaron	Mitchell	James Brown	19.799999999999997
4	François	Tremblay	The Who	19.799999999999997
5	Robert	Brown	Creedence Clearwater Revival	19.799999999999997
6	Helena	Holý	Red Hot Chili Peppers	19.799999999999997
7	R	Madhav	Kiss	19.799999999999997
8	Heather	Leacock	House Of Pain	18.81
9	Niklas	Schröder	Queen	18.81
10	Hugh	O'Reilly	Nirvana	18.81
11	Hugh	O'Reilly	Marisa Monte	17.82
12	Camille	Bernard	Marisa Monte	17.82
13	Luís	Gonçalves	The Cult	17.82
14	Mark	Taylor	The Clash	17.82
15	Steve	Murray	AC/DC	17.82
16	Richard	Cunningham	Marvin Gaye	17.82

Total rows: 1000 of 2189 | Query complete 00:00:00.260 | Ln 125, Col 16

/* 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. */

/***Steps to solve**:In the final table music genre ,country and heighest purchases of top genre album

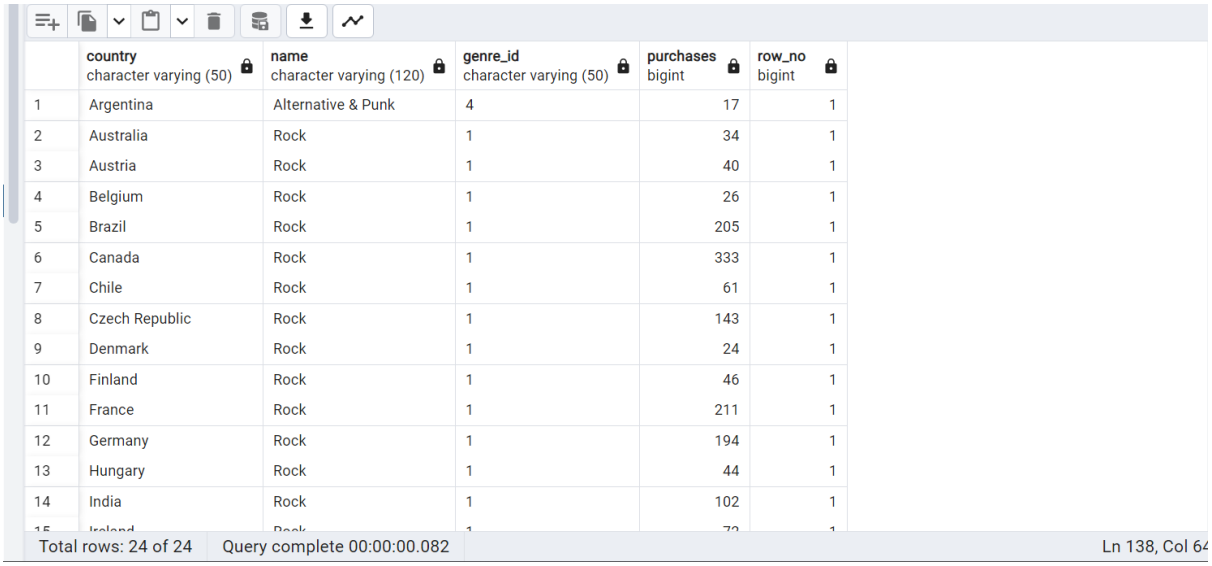
is needed.So customer.country,genre.name,genre.genre_id,count(invoice_line.quantity) as purchases

has been slected.It is asked top genre in each country so windows function ROW_NUMBER()is used.And then storing all the columns in a temporary table named " Highest_amount_purchases " using CTE.From this table it can be seen all the different genre that has been purchased in different countries. So to get the top genre I have applied the crieteria Row_no<=1 using where clause */

Query:

```
with Highest_amount_purchases as(
select customer.country,genre.name,genre.genre_id,count(invoice_line.quantity) as purchases,
 ROW_NUMBER() over(partition by customer.country order by count(invoice_line.quantity) DESC)as Row_no
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 track.track_id=invoice_line.track_id
join genre on track.genre_id=genre.genre_id
group by 1,2,3
order by 1 asc,4 desc
)
select * from Highest_amount_purchases
where Row_no<=1
```

O/P



	country character varying (50)	name character varying (120)	genre_id character varying (50)	purchases bigint	row_no bigint
1	Argentina	Alternative & Punk	4	17	1
2	Australia	Rock	1	34	1
3	Austria	Rock	1	40	1
4	Belgium	Rock	1	26	1
5	Brazil	Rock	1	205	1
6	Canada	Rock	1	333	1
7	Chile	Rock	1	61	1
8	Czech Republic	Rock	1	143	1
9	Denmark	Rock	1	24	1
10	Finland	Rock	1	46	1
11	France	Rock	1	211	1
12	Germany	Rock	1	194	1
13	Hungary	Rock	1	44	1
14	India	Rock	1	102	1
15	Ireland	Rock	1	72	1

Total rows: 24 of 24 Query complete 00:00:00.082 Ln 138, Col 64

/* 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. */

/***Steps to solve**:In the final output I need country name,top customer and their spent amount.

for that I have selected customer_id,first_name,last_name from "customer" table and "Total" from Invoice

table.For accumlating the mentioned columns I have to join customer with invoice on customer_id.

Now as the customer with top amount spent has asked I have used ROW_NO() and later on used Row_no<=1

to get the top amount spent customer. */

QUERY

```
with Top_amount_spent as(
select customer.customer_id,customer.first_name,customer.last_name,invoice.billing_country,
sum(invoice.total) as Total_spending,
ROW_NUMBER() over(partition by customer.country order by sum(invoice.total)DESC)as Row_no
from customer
join invoice on customer.customer_id=invoice.customer_id
group by 1,2,3,4
order by 4 asc,5 desc
)
select * from Top_amount_spent
where Row_no<=1
```

	customer_id integer	first_name character	last_name character	billing_country character varying (30)	total_spending double precision	row_no bigint
1	56	Diego	Gutiérrez	Argentina	39.6	1
2	55	Mark	Taylor	Australia	81.18	1
3	7	Astrid	Gruber	Austria	69.3	1
4	8	Daan	Peeters	Belgium	60.38999999999999	1
5	1	Luís	Gonçalves	Brazil	108.89999999999998	1
6	3	François	Tremblay	Canada	99.99	1
7	57	Luis	Rojas	Chile	97.02000000000001	1
8	5	R	Madhav	Czech Republic	144.54000000000002	1
9	9	Kara	Nielsen	Denmark	37.61999999999999	1
10	44	Terhi	Hämäläinen	Finland	79.2	1
11	42	Wyatt	Girard	France	99.99	1
12	37	Fynn	Zimmermann	Germany	94.05000000000001	1
13	45	Ladislav	Kovács	Hungary	78.21	1
14	58	Manoj	Pareek	India	111.86999999999999	1
15	46	Hugh	O'Reilly	Ireland	114.83999999999997	1

Total rows: 24 of 24 Query complete 00:00:00.085 Ln 177, Col 16