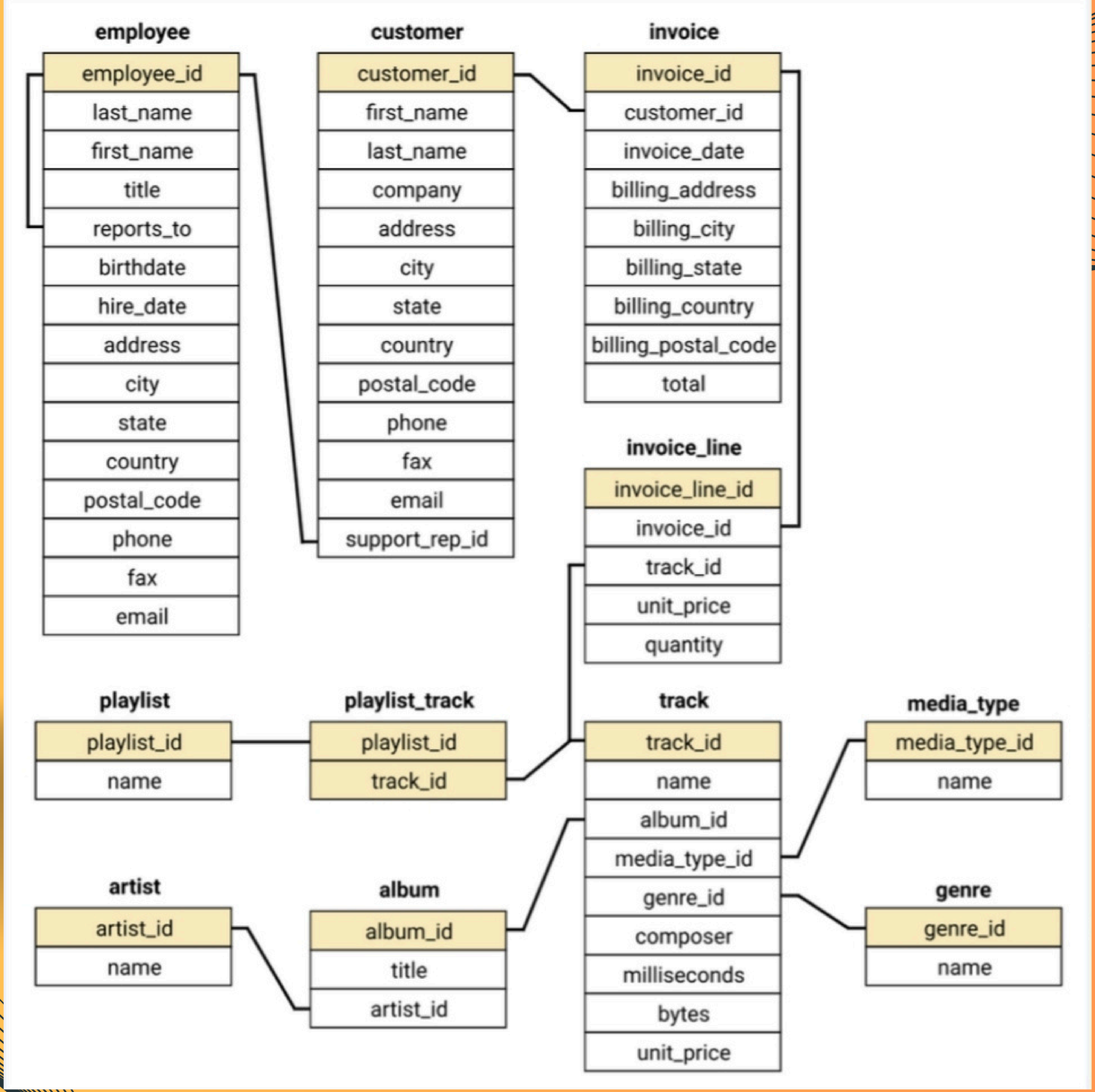
Music Store Analysis - SQL CASE STUDY

By- Aman Sharma

Schema Diagram-



employee
- employee_id
- last_name
- first_name
- title
- reports_to
- birthdate
- hire_date
- address
- city
- state
- country
- postal_code
- phone
- fax
- email

customer
- customer_id
- first_name
- last_name
- company
- address
- city
- state
- country
- postal_code
- phone
- fax
- email
- support_rep_id

invoice
- invoice_id
- customer_id
- invoice_date
- billing_address
- billing_city
- billing_state
- billing_country
- billing_postal_code
- total

invoice_line
- invoice_line_id
- invoice_id
- track_id
- unit_price
- quantity

playlist
- playlist_id
- name

playlist_track
- playlist_id
- track_id

track
- track_id
- name
- album_id
- media_type_id
- genre_id
- composer
- milliseconds
- bytes
- unit_price

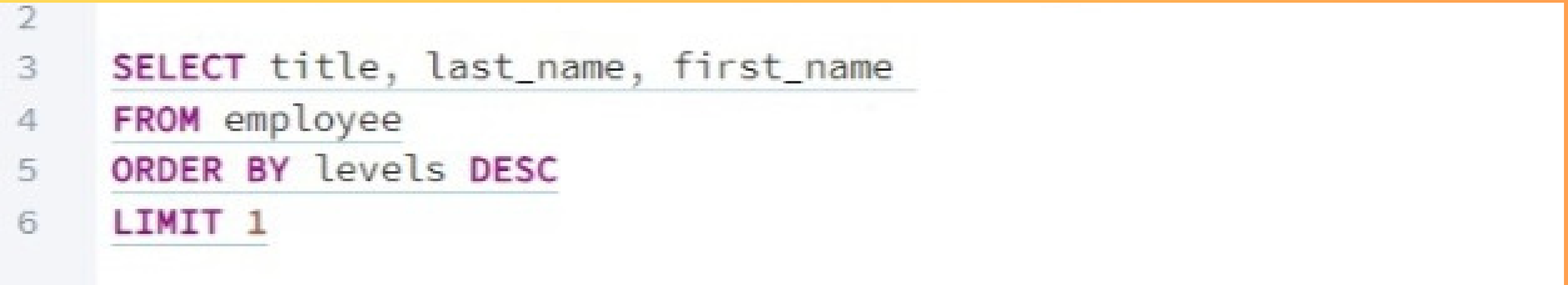
media_type
- media_type_id
- name

artist
- artist_id
- name

album
- album_id
- title
- artist_id

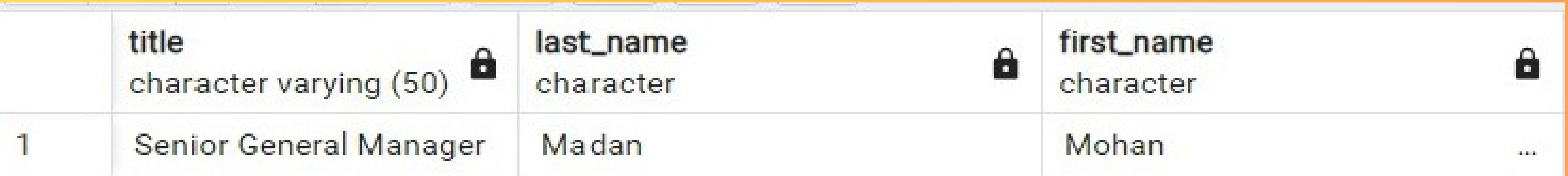
genre
- genre_id
- name

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



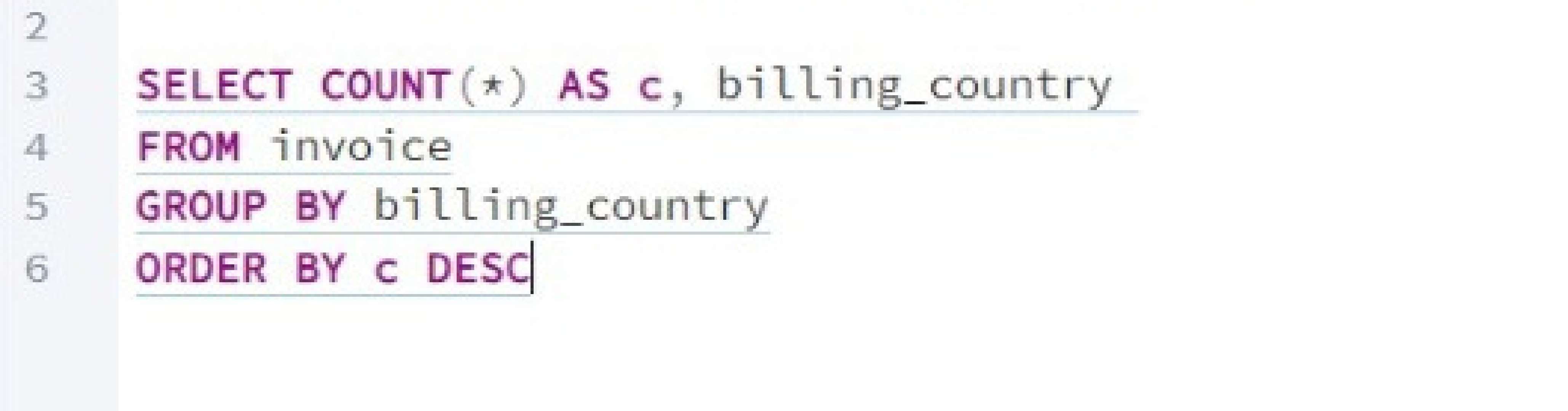
```sql

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

OUTPUT:-



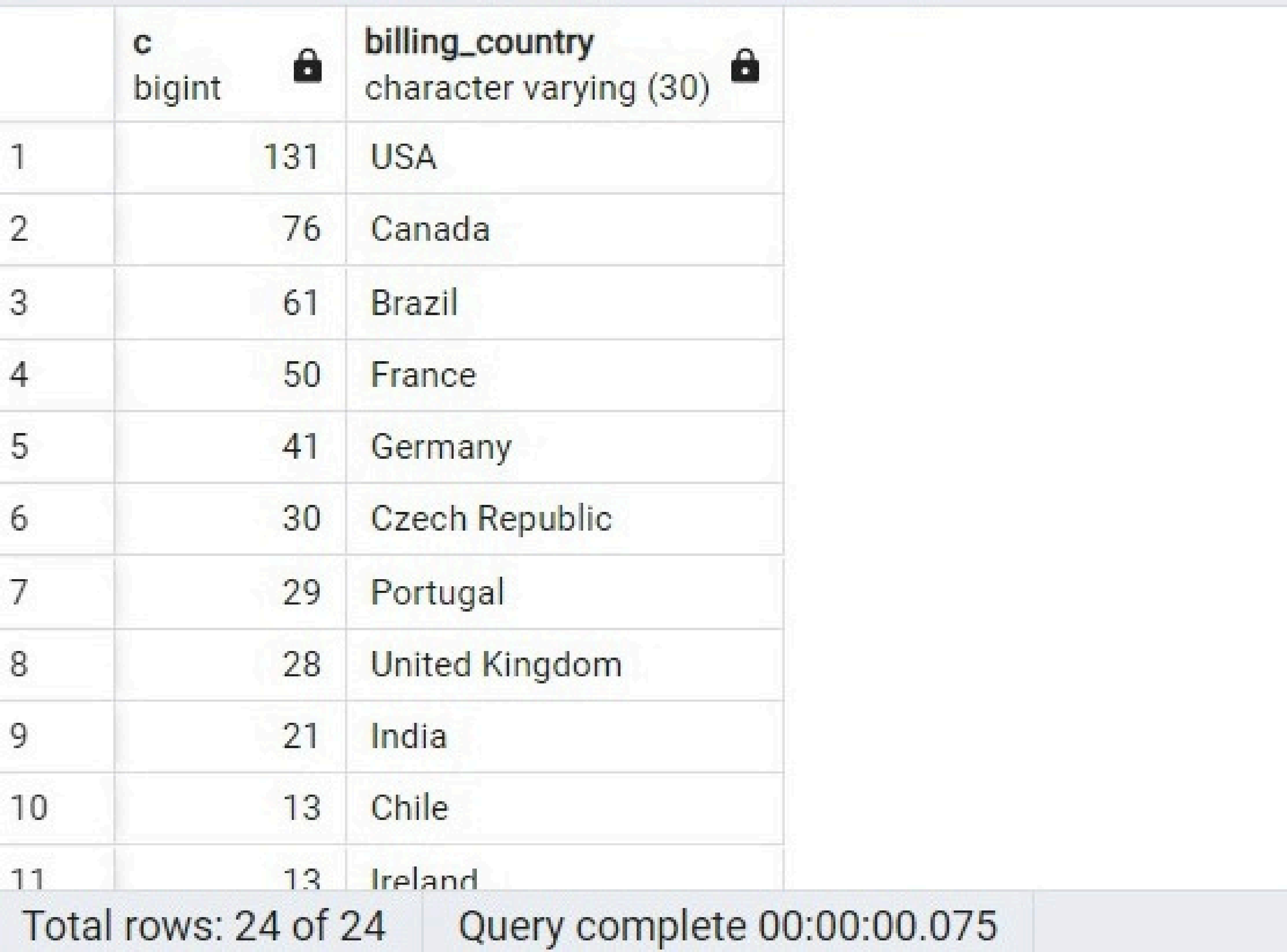
	title character varying (50)	last_name character	first_name character
1	Senior General Manager	Madan	Mohan

2.Which countries have the most invoices?



```sql

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

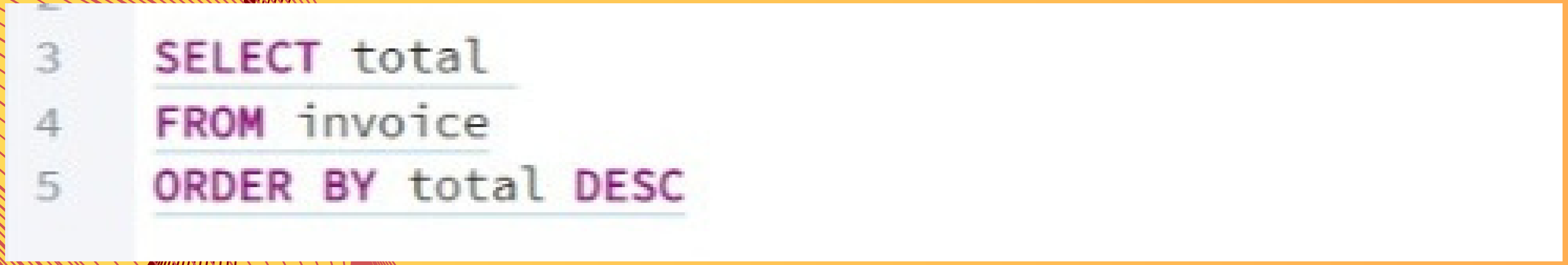
OUTPUT:-



	c bigint	billing_country character varying (30)
1	131	USA
2	76	Canada
3	61	Brazil
4	50	France
5	41	Germany
6	30	Czech Republic
7	29	Portugal
8	28	United Kingdom
9	21	India
10	13	Chile
11	13	Ireland

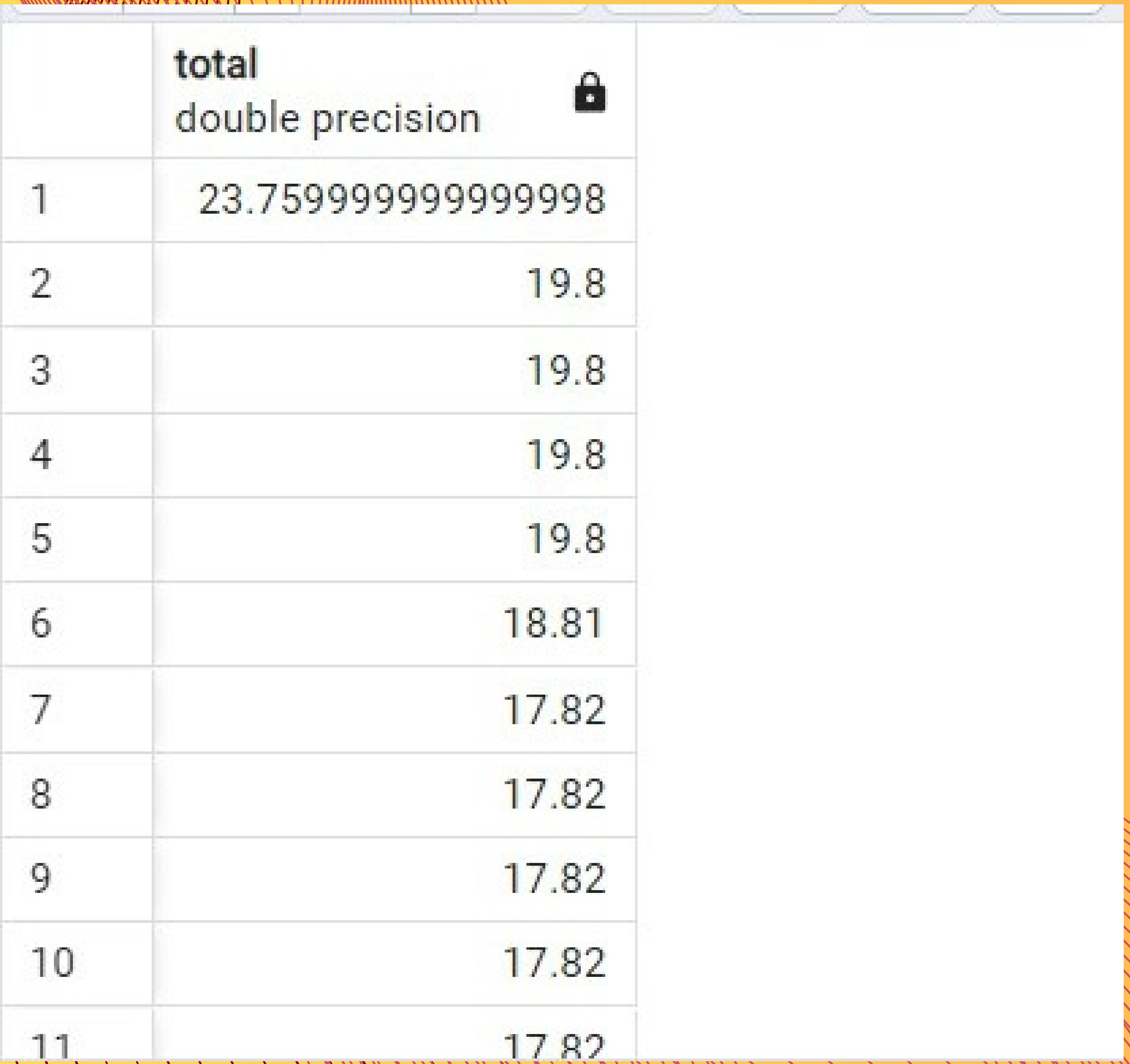
Total rows: 24 of 24 Query complete 00:00:00.075

3. What are top 3 values of total invoice?



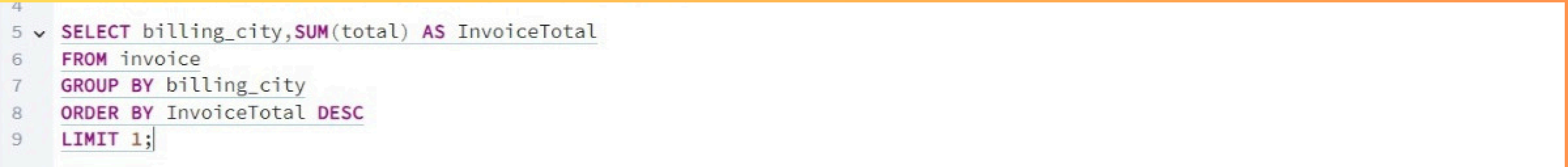
```
SELECT total
FROM invoice
ORDER BY total DESC
```

OUTPUT:-



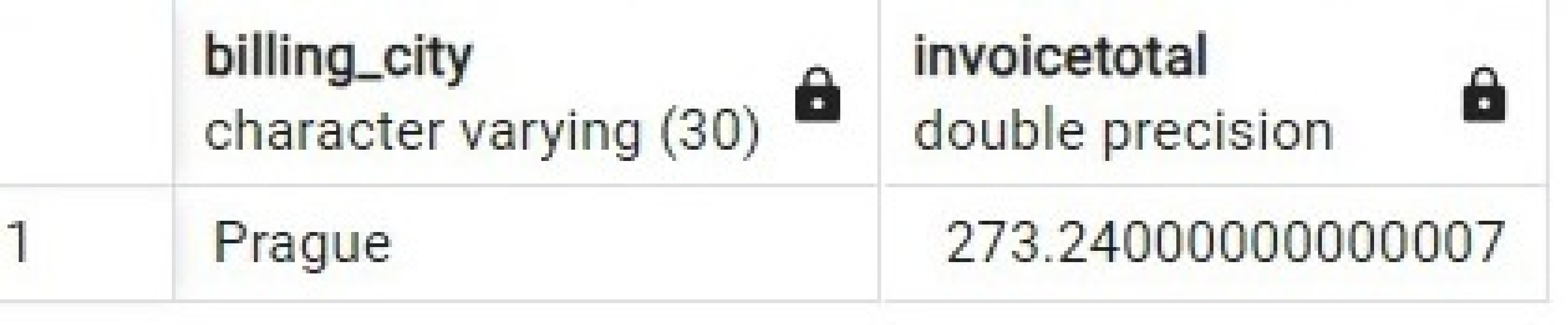
	total double precision
1	23.759999999999998
2	19.8
3	19.8
4	19.8
5	19.8
6	18.81
7	17.82
8	17.82
9	17.82
10	17.82
11	17.82

4. Which city has the best customers?



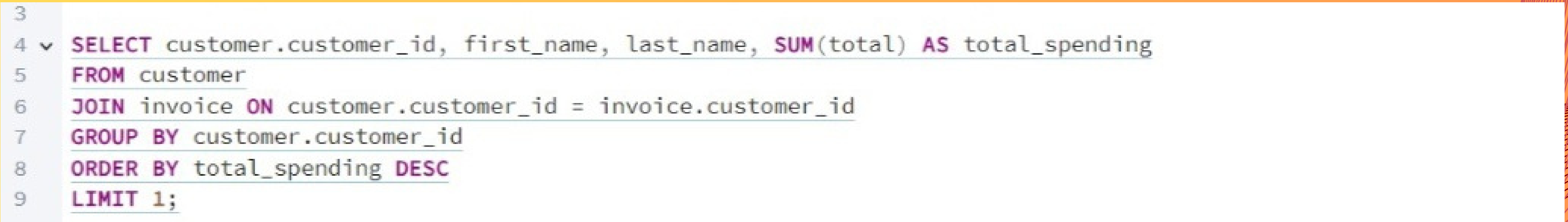
```
SELECT billing_city,SUM(total) AS InvoiceTotal
FROM invoice
GROUP BY billing_city
ORDER BY InvoiceTotal DESC
LIMIT 1;
```

OUTPUT:-



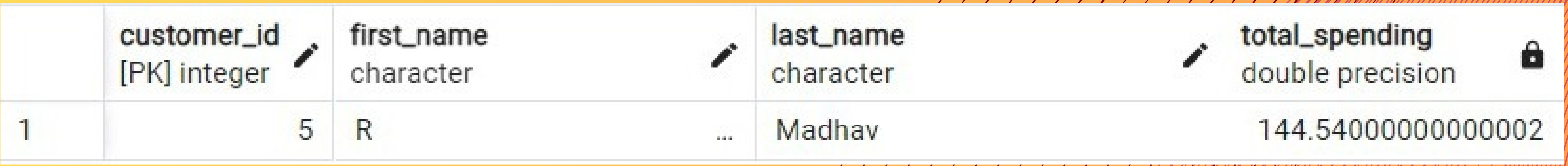
	billing_city character varying (30)	invoicetotal double precision
1	Prague	273.24000000000007

5. Who is the best customer?



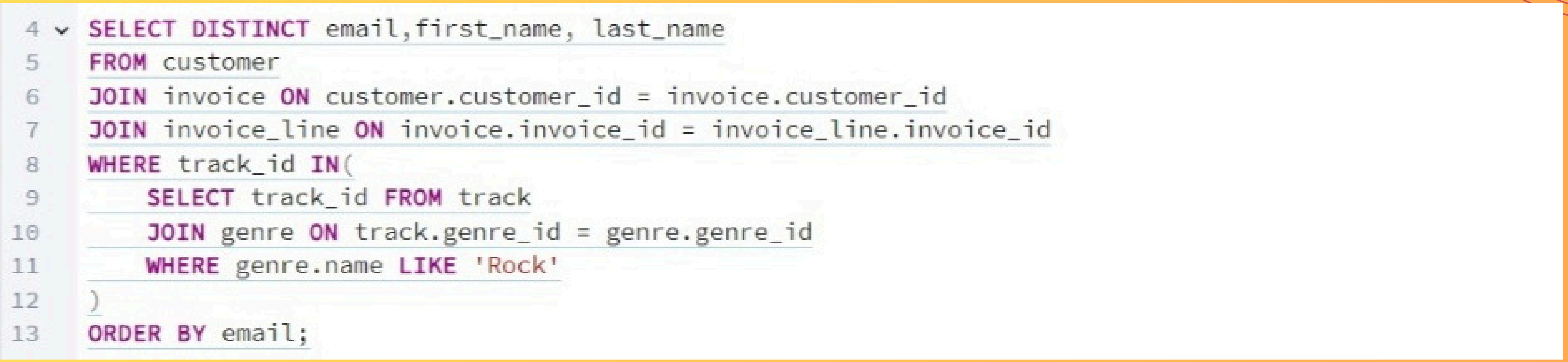
```sql
SELECT customer.customer_id, first_name, last_name, SUM(total) AS total_spending
FROM customer
JOIN invoice ON customer.customer_id = invoice.customer_id
GROUP BY customer.customer_id
ORDER BY total_spending DESC
LIMIT 1;
```

OUTPUT:-



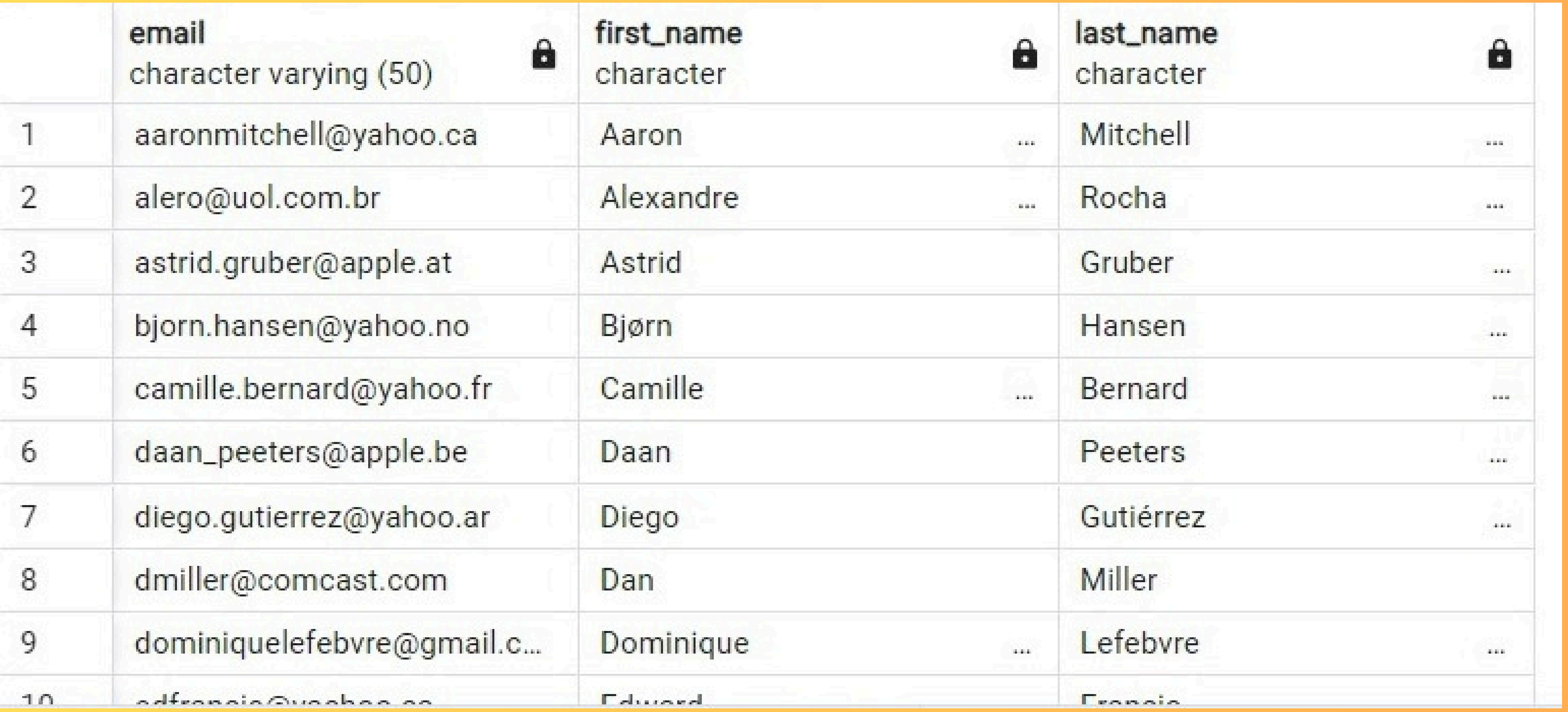
	customer_id	first_name	last_name	total_spending
	[PK] integer	character	character	double precision
1	5	R	Madhav	144.54000000000002

6. Write query to return the email, First name, Last name,& genre of all rocks music listeners. Return your list ordered alphabetically by email starting with A ?



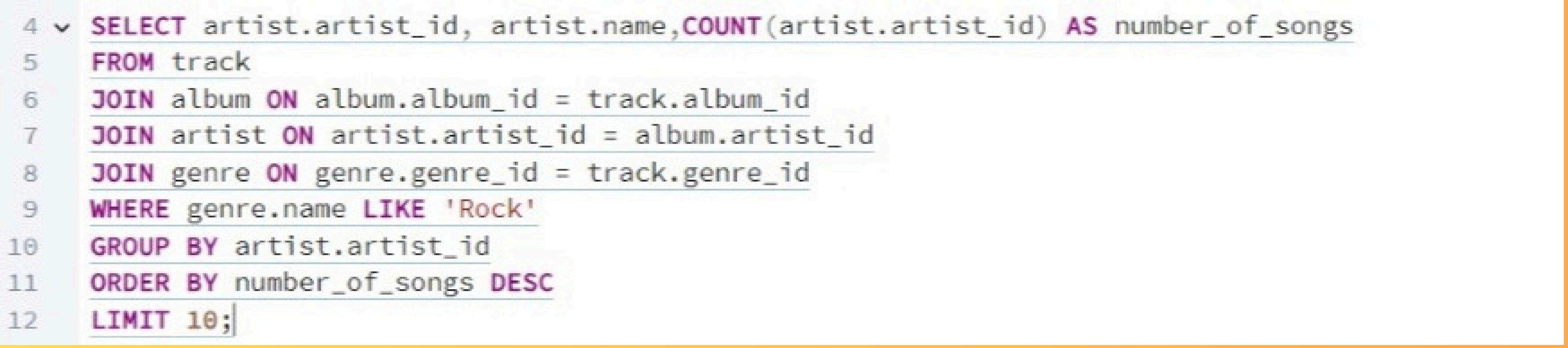
```sql
SELECT DISTINCT email,first_name, last_name
FROM customer
JOIN invoice ON customer.customer_id = invoice.customer_id
JOIN invoice_line ON invoice.invoice_id = invoice_line.invoice_id
WHERE track_id IN(
    SELECT track_id FROM track
    JOIN genre ON track.genre_id = genre.genre_id
    WHERE genre.name LIKE 'Rock'
)
ORDER BY email;
```

OUTPUT:-



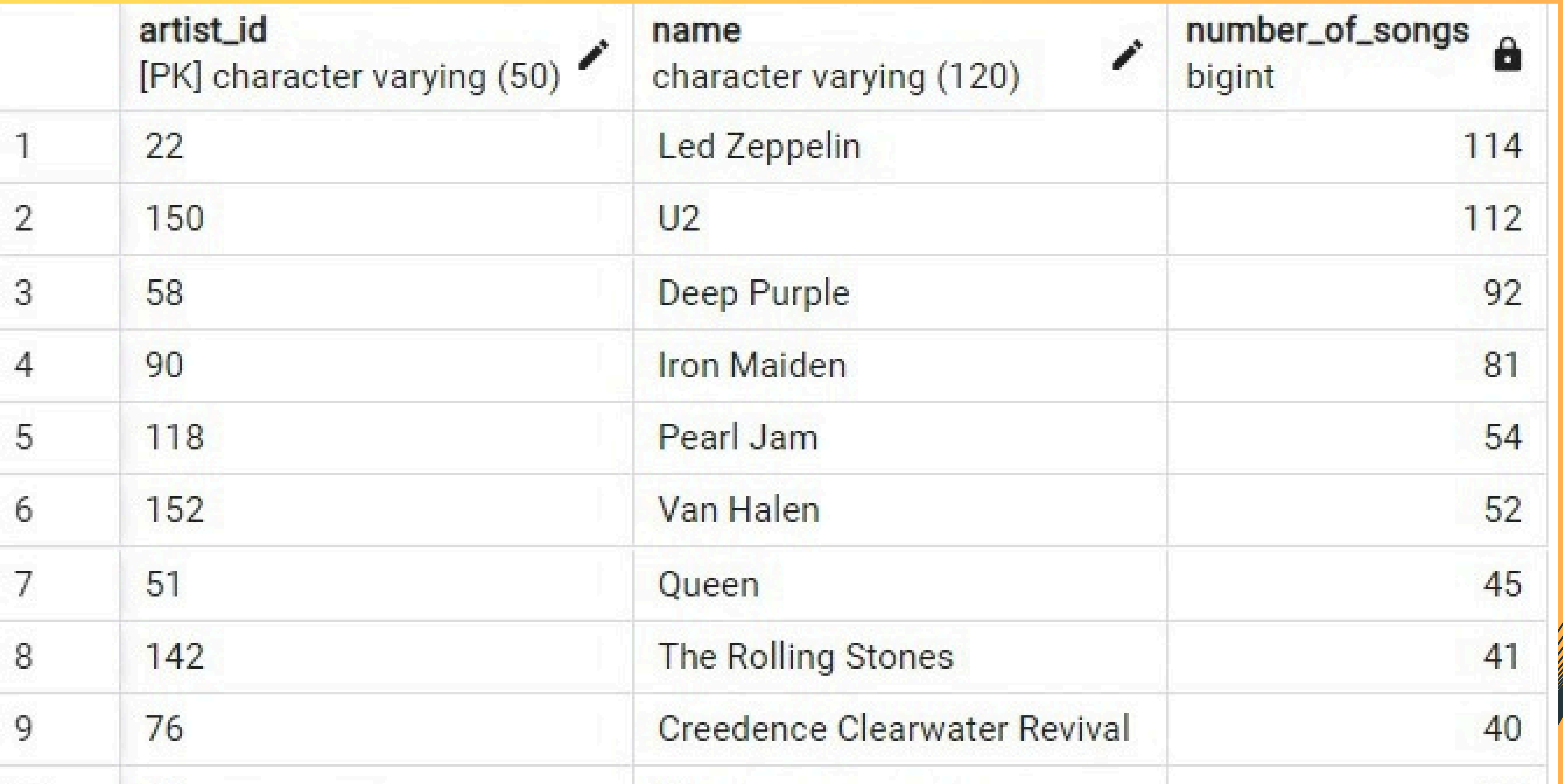
	email character varying (50)	first_name character	last_name character
1	aaronmitchell@yahoo.ca	Aaron ...	Mitchell ...
2	alero@uol.com.br	Alexandre ...	Rocha ...
3	astrid.gruber@apple.at	Astrid	Gruber ...
4	bjorn.hansen@yahoo.no	Bjørn	Hansen ...
5	camille.bernard@yahoo.fr	Camille ...	Bernard ...
6	daan_peeters@apple.be	Daan	Peeters ...
7	diego.gutierrez@yahoo.ar	Diego	Gutiérrez ...
8	dmiller@comcast.com	Dan	Miller
9	dominiquelefebvre@gmail.c...	Dominique ...	Lefebvre ...
10	edfrancis@yahoo.ca	Edward	Francis

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



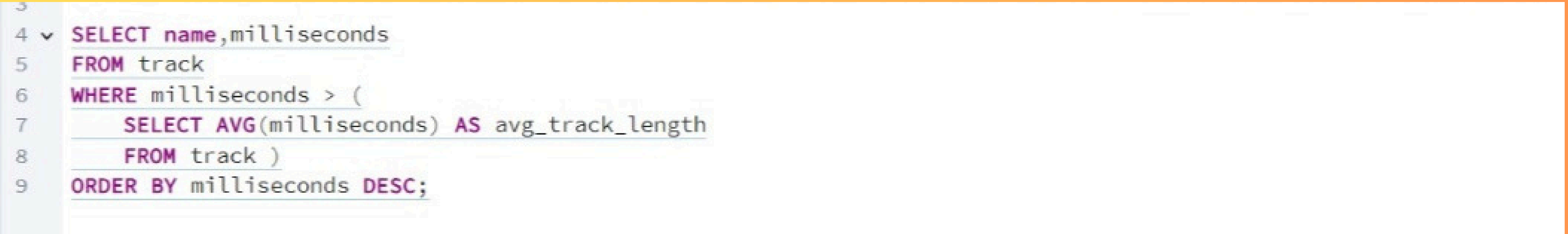
```sql
SELECT artist.artist_id, artist.name,COUNT(artist.artist_id) AS number_of_songs
FROM track
JOIN album ON album.album_id = track.album_id
JOIN artist ON artist.artist_id = album.artist_id
JOIN genre ON genre.genre_id = track.genre_id
WHERE genre.name LIKE 'Rock'
GROUP BY artist.artist_id
ORDER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

OUTPUT:-



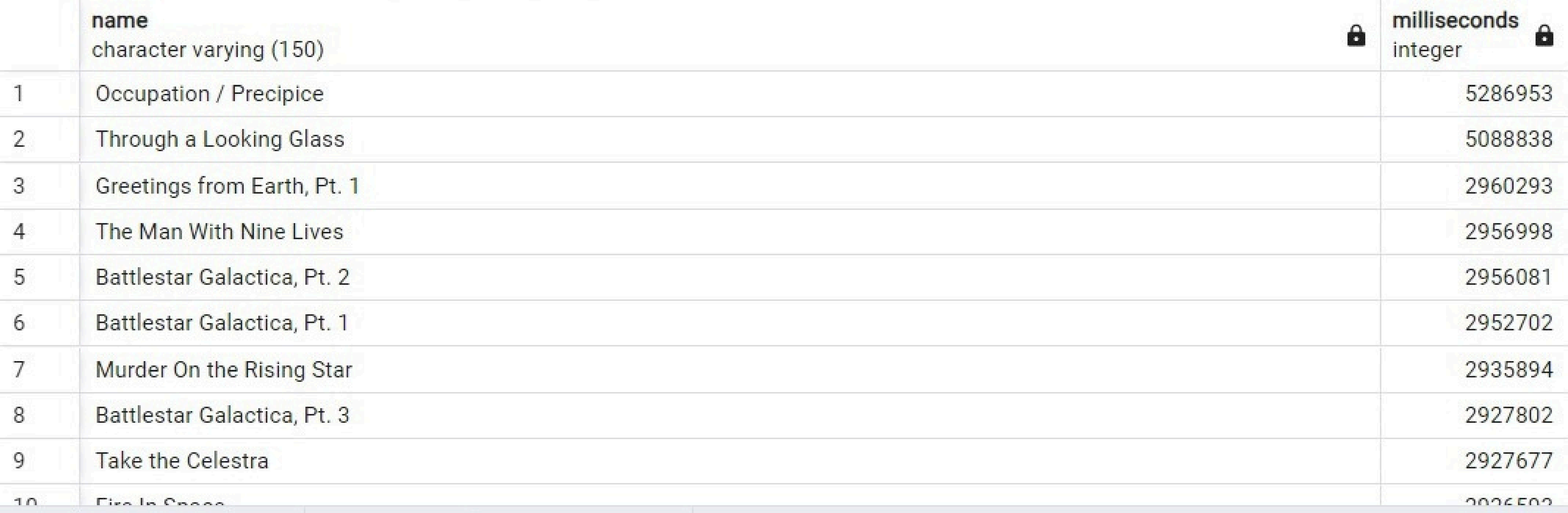
	artist_id [PK] character varying (50)	name character varying (120)	number_of_songs bigint
1	22	Led Zeppelin	114
2	150	U2	112
3	58	Deep Purple	92
4	90	Iron Maiden	81
5	118	Pearl Jam	54
6	152	Van Halen	52
7	51	Queen	45
8	142	The Rolling Stones	41
9	76	Creedence Clearwater Revival	40

8. Return all the track names that have a song length longer than the average song length. Return the names and milliseconds for each track. Longest songs listed first.



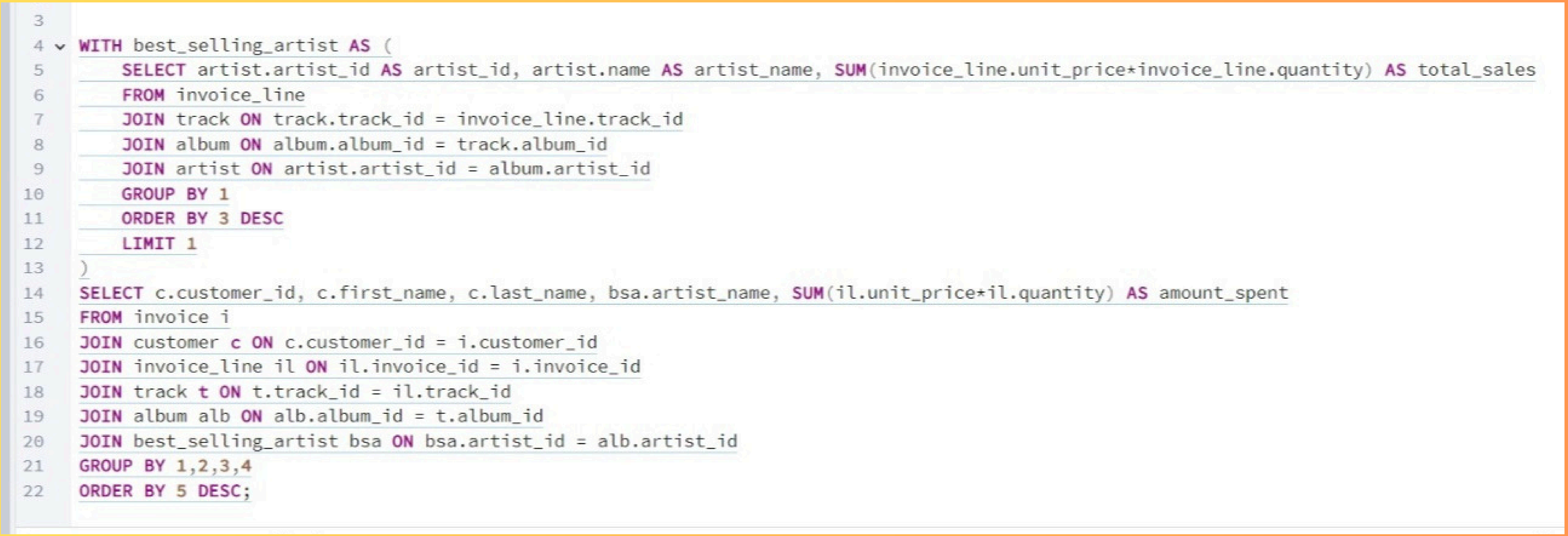
```
SELECT name,milliseconds
FROM track
WHERE milliseconds > (
    SELECT AVG(milliseconds) AS avg_track_length
    FROM track )
ORDER BY milliseconds DESC;
```

OUTPUT:-



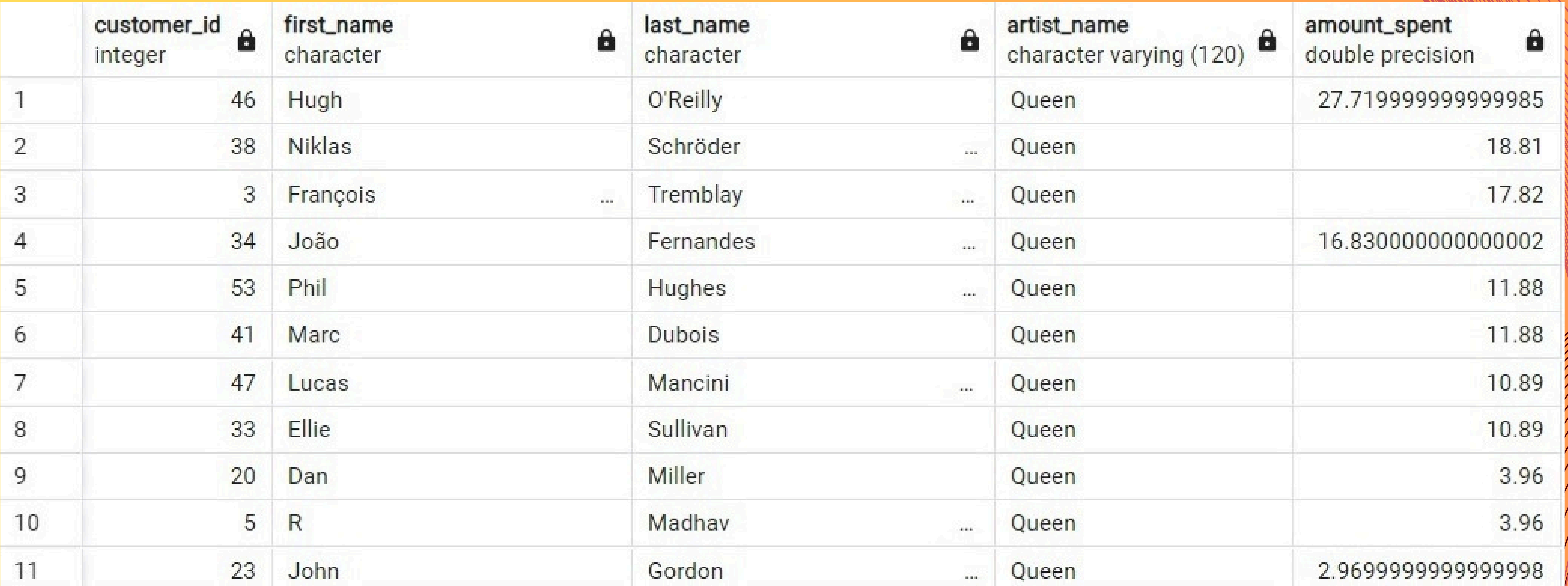
	name character varying (150)	milliseconds integer
1	Occupation / Precipice	5286953
2	Through a Looking Glass	5088838
3	Greetings from Earth, Pt. 1	2960293
4	The Man With Nine Lives	2956998
5	Battlestar Galactica, Pt. 2	2956081
6	Battlestar Galactica, Pt. 1	2952702
7	Murder On the Rising Star	2935894
8	Battlestar Galactica, Pt. 3	2927802
9	Take the Celestra	2927677
10	Fire In Space	2926593

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



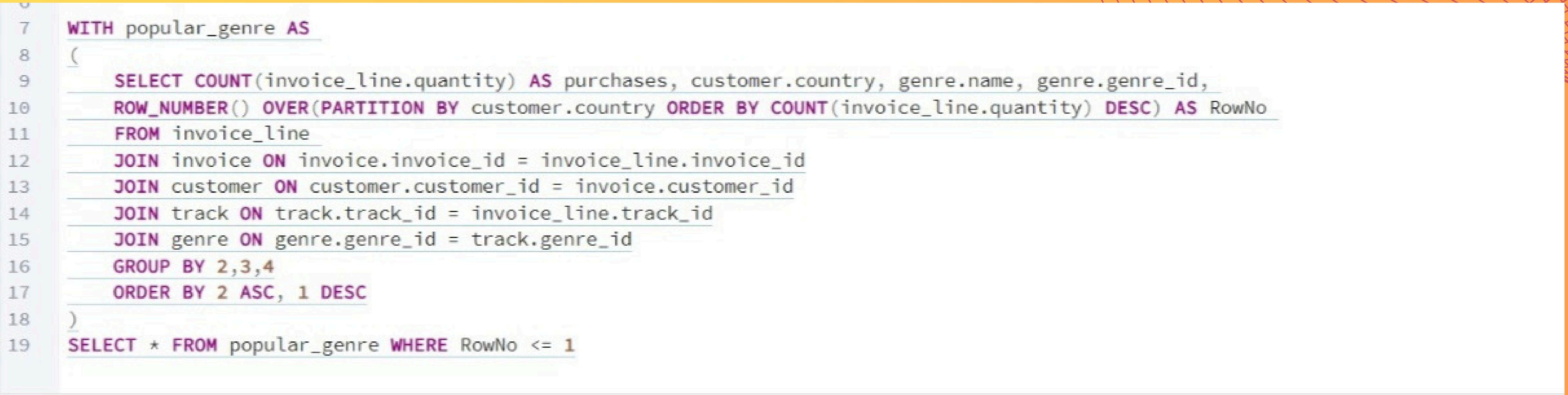
```sql
WITH best_selling_artist AS (
    SELECT artist.artist_id AS artist_id, artist.name AS artist_name, SUM(invoice_line.unit_price*invoice_line.quantity) AS total_sales
    FROM invoice_line
    JOIN track ON track.track_id = invoice_line.track_id
    JOIN album ON album.album_id = track.album_id
    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;
```

OUTPUT:-



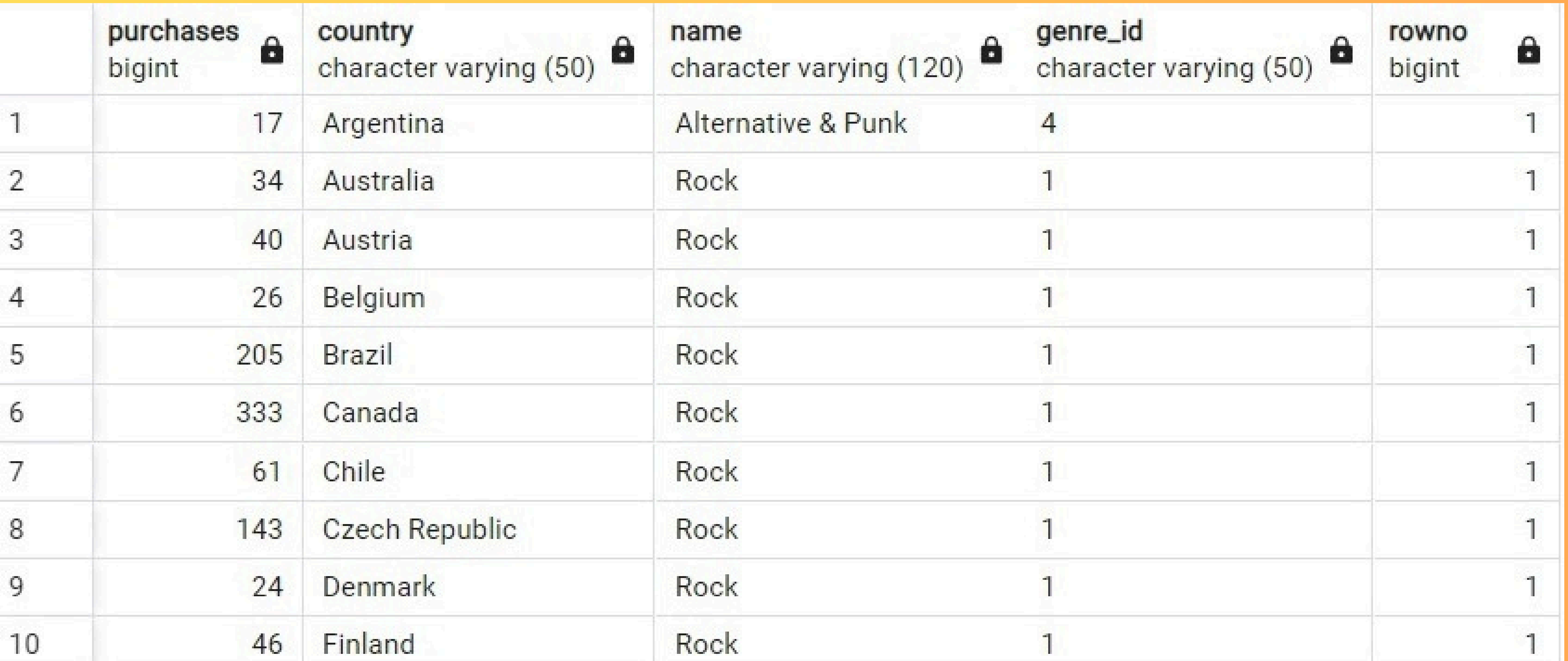
	customer_id	first_name	last_name	artist_name	amount_spent
	integer	character	character	character varying (120)	double precision
1	46	Hugh	O'Reilly	Queen	27.719999999999985
2	38	Niklas	Schröder ...	Queen	18.81
3	3	François ...	Tremblay ...	Queen	17.82
4	34	João	Fernandes ...	Queen	16.830000000000002
5	53	Phil	Hughes ...	Queen	11.88
6	41	Marc	Dubois	Queen	11.88
7	47	Lucas	Mancini ...	Queen	10.89
8	33	Ellie	Sullivan	Queen	10.89
9	20	Dan	Miller	Queen	3.96
10	5	R	Madhav ...	Queen	3.96
11	23	John	Gordon ...	Queen	2.9699999999999998

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 purchase. Write a query that returns each country along with top genre.



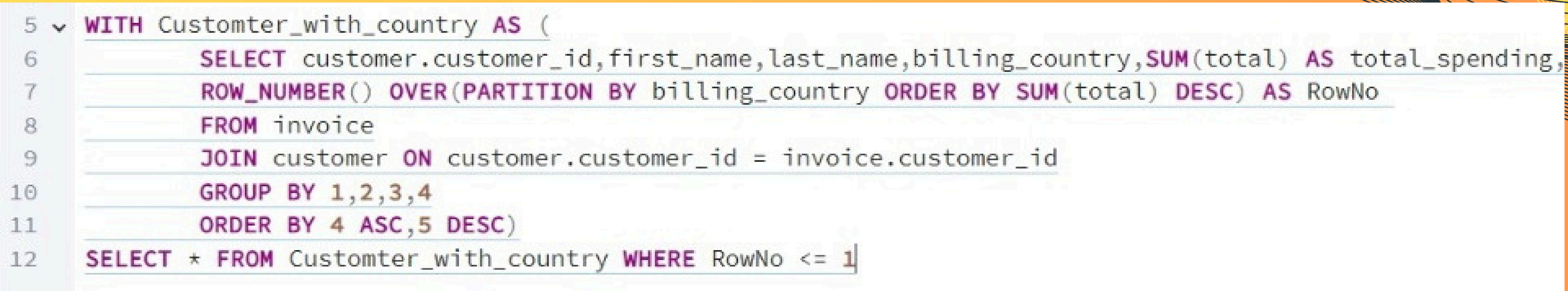
```sql
WITH popular_genre AS
(
    SELECT COUNT(invoice_line.quantity) AS purchases, customer.country, genre.name, genre.genre_id,
    ROW_NUMBER() OVER(PARTITION BY customer.country ORDER BY COUNT(invoice_line.quantity) DESC) AS RowNo
    FROM invoice_line
    JOIN invoice ON invoice.invoice_id = invoice_line.invoice_id
    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
```

OUTPUT:-



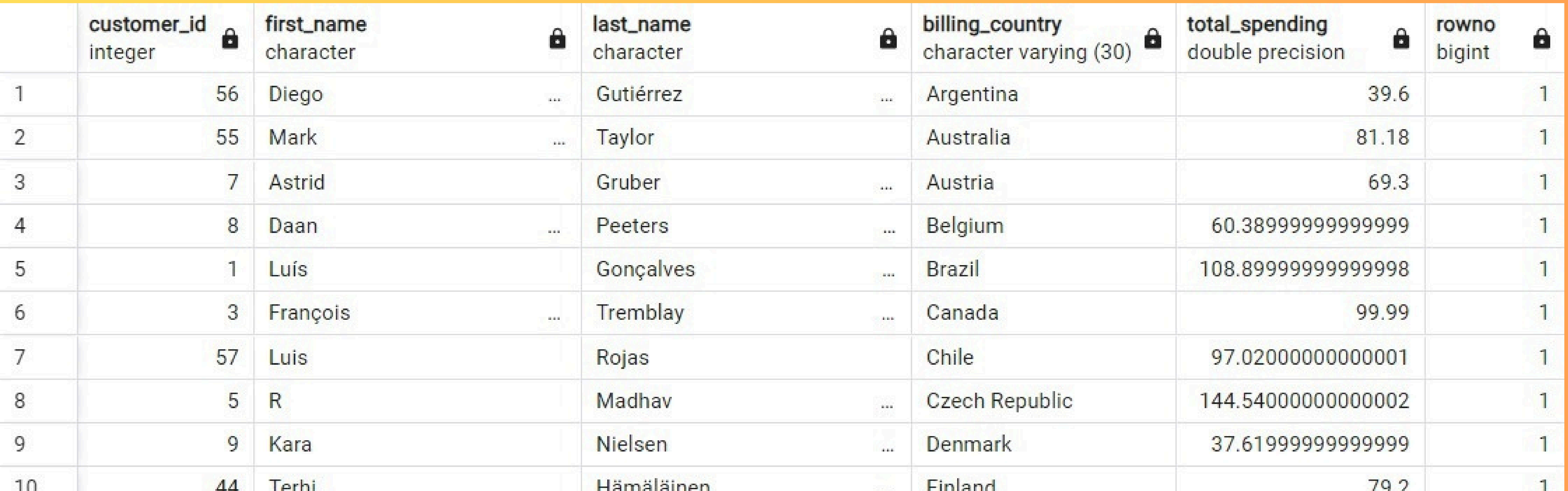
	purchases bigint	country character varying (50)	name character varying (120)	genre_id character varying (50)	rowno bigint
1	17	Argentina	Alternative & Punk	4	1
2	34	Australia	Rock	1	1
3	40	Austria	Rock	1	1
4	26	Belgium	Rock	1	1
5	205	Brazil	Rock	1	1
6	333	Canada	Rock	1	1
7	61	Chile	Rock	1	1
8	143	Czech Republic	Rock	1	1
9	24	Denmark	Rock	1	1
10	46	Finland	Rock	1	1

11. Write a query that determine the customer that has spet the most on music for each country. Write a query that returns the country along with the top customer and how much they spent.



```
WITH Customter_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 Customter_with_country WHERE RowNo <= 1
```

OUTPUT:-



	customer_id integer	first_name character	last_name character	billing_country character varying (30)	total_spending double precision	rowno 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

THANK YOU

aamansharma027@gmail.com ✉