SQL PROJECT - Digital Music Store Analysis

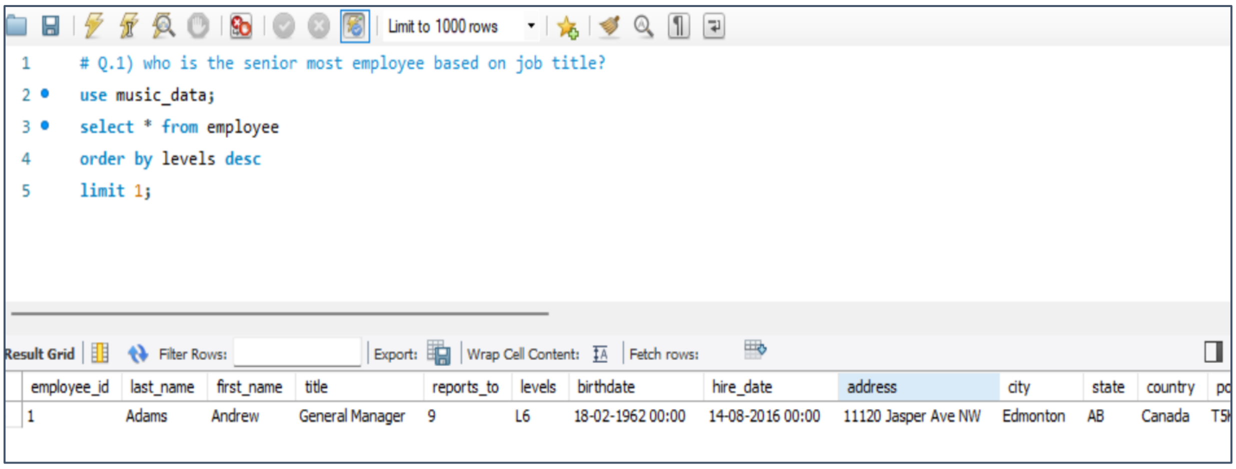
<u>Objective</u> –

- To analyse the music playlist database.
- Examine the dataset with SQL and help the store understand its business growth.

<u>Learning</u> -

- Focused on understanding the data schema, specifically identifying the tables and their relationships, including the complexities within each table.
- Checked all missing values, null columns ensuring proper text and validating headers in the dataset.
- Joins, different analytical and aggregate functions, group by, window function, CTE were using to get the proper insights.

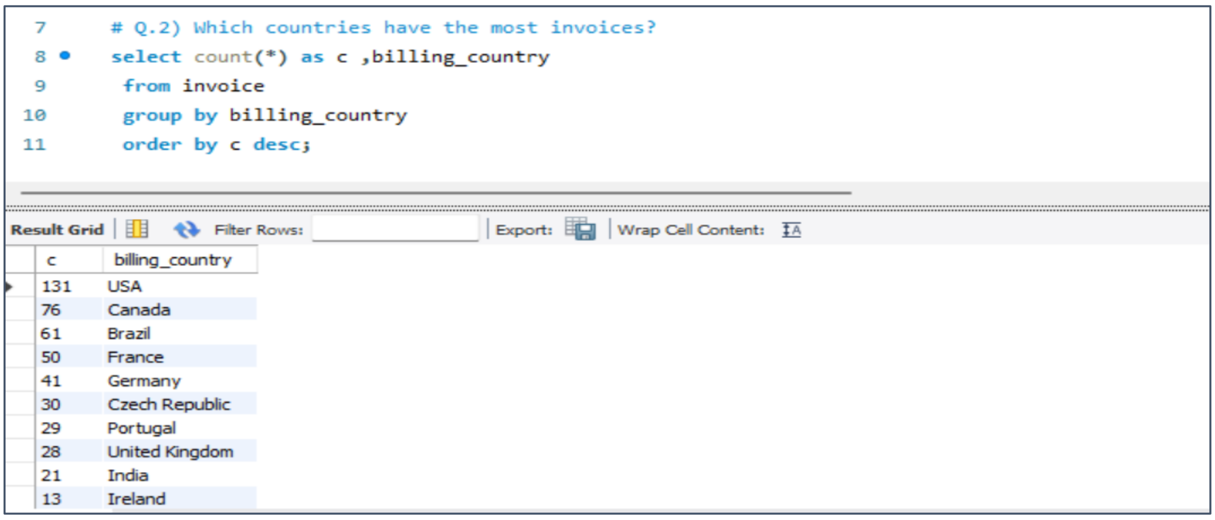
- 1)Who is the senior most employee based in job title?



```
# Q.1) who is the senior most employee based on job title?
use music_data;
select * from employee
order by levels desc
limit 1;
```

```
employee_id  last_name  first_name  title            reports_to  levels  birthdate         hire_date         address              city      state  country  po
1            Adams      Andrew      General Manager  9           L6      18-02-1962 00:00  14-08-2016 00:00  11120 Jasper Ave NW  Edmonton  AB     Canada   T5
```

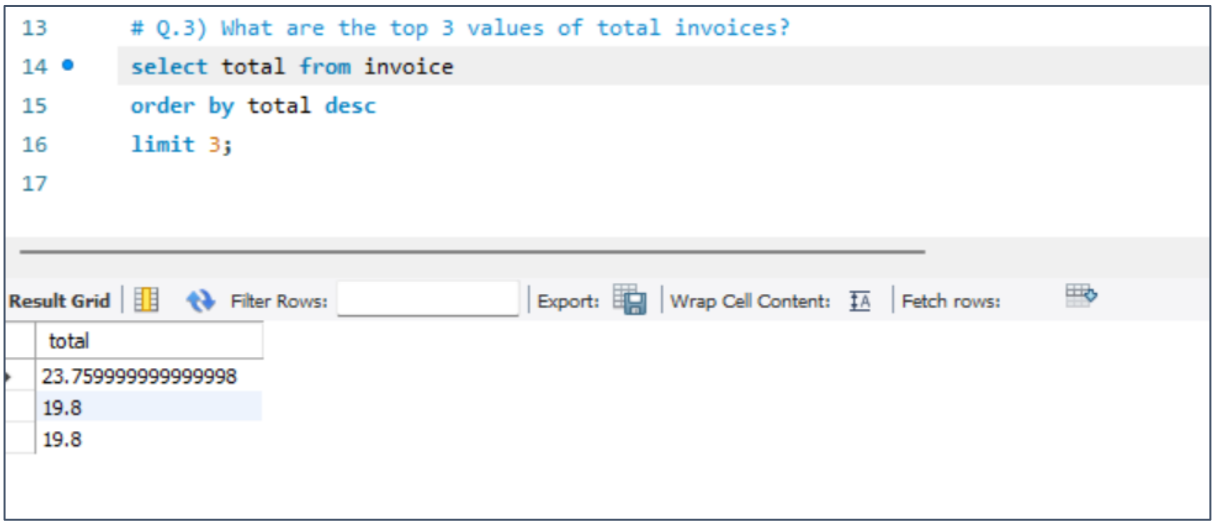
- **2)Which countries have the most invoices?**



```
# Q.2) Which countries have the most invoices?
select count(*) as c ,billing_country
 from invoice
 group by billing_country
 order by c desc;
```

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

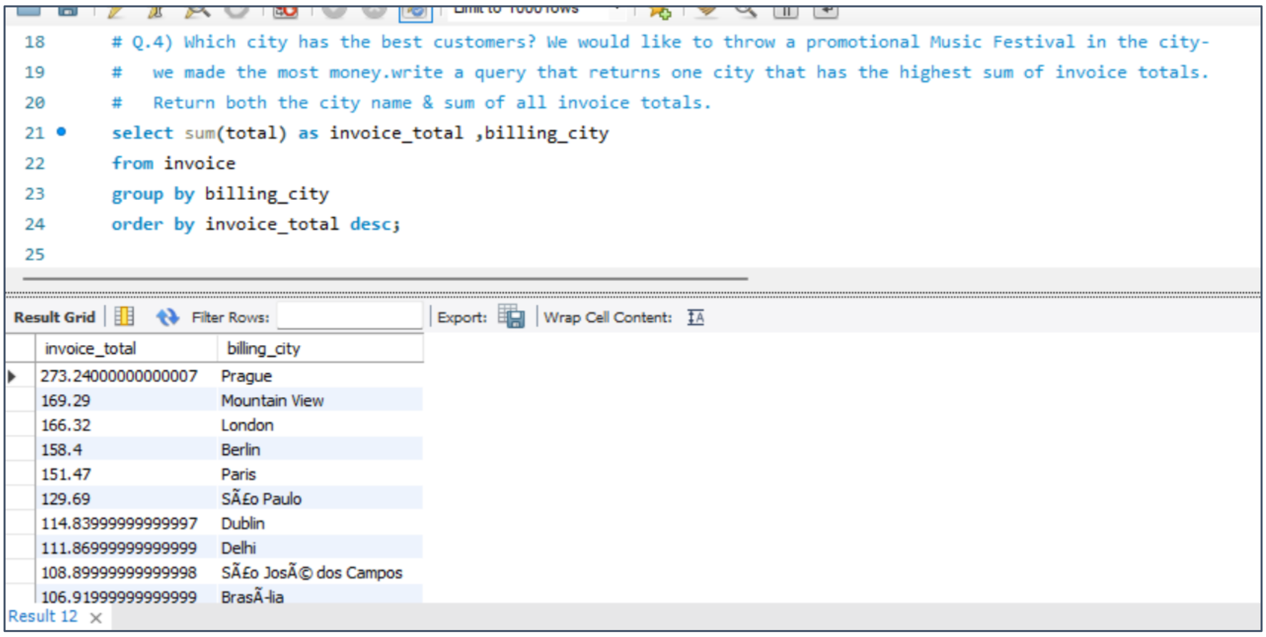
- 3)What are the top 3 values of total invoices?



```
# Q.3) What are the top 3 values of total invoices?
select total from invoice
order by total desc
limit 3;
```

```
total
23.759999999999998
19.8
19.8
```

- 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.

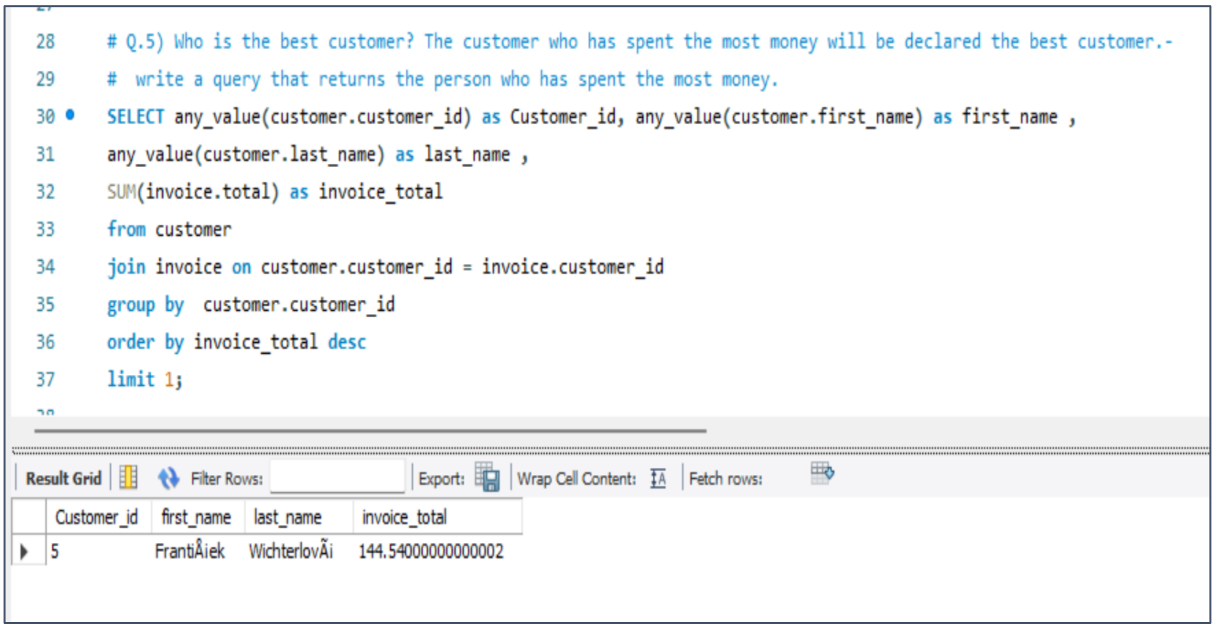


```
# Q.4) Which city has the best customers? We would like to throw a promotional Music Festival in the city-
#  we made the most money.write a query that returns one city that has the highest sum of invoice totals.
#  Return both the city name & sum of all invoice totals.
select sum(total) as invoice_total ,billing_city
from invoice
group by billing_city
order by invoice_total desc;
```

invoice_total	billing_city
273.24000000000007	Prague
169.29	Mountain View
166.32	London
158.4	Berlin
151.47	Paris
129.69	SÃ£o Paulo
114.83999999999997	Dublin
111.86999999999999	Delhi
108.89999999999998	SÃ£o JosÃ© dos Campos
106.91999999999999	BrasÃ-lia

Result 12

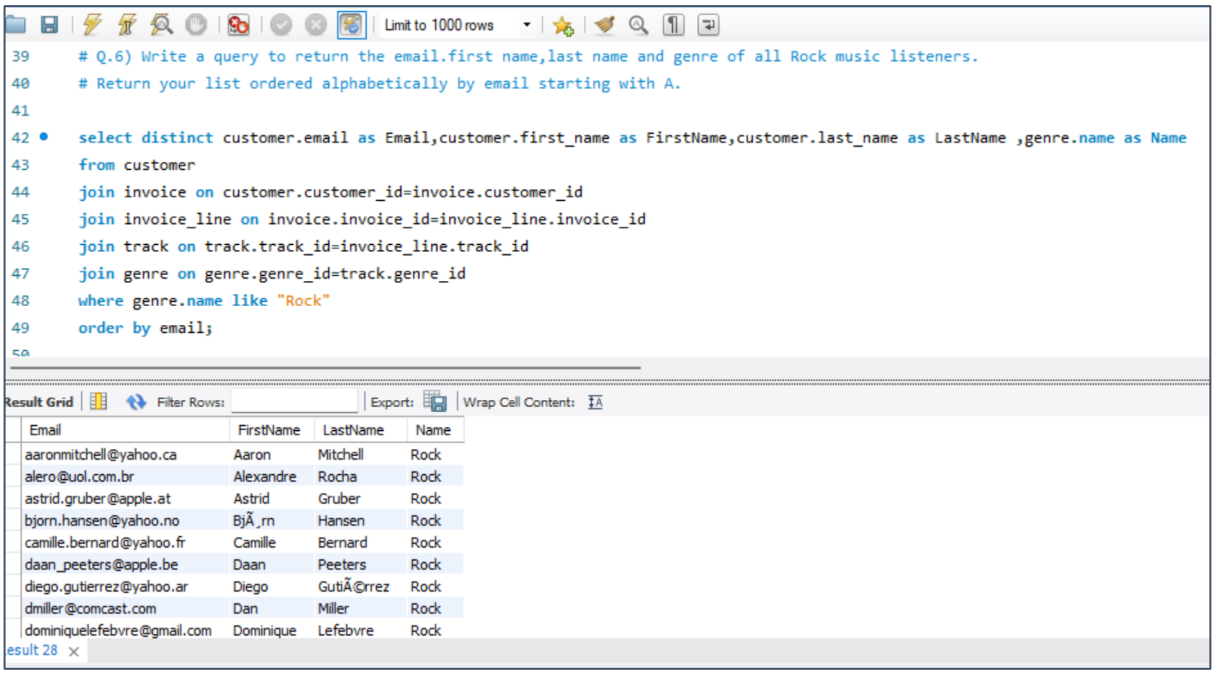
- 5) Who is the best customer? The customer who has spent the most money will be declared the best customer. write a query that returns the person who has spent the most money.



```
# Q.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 any_value(customer.customer_id) as Customer_id, any_value(customer.first_name) as first_name ,
any_value(customer.last_name) as last_name ,
SUM(invoice.total) as invoice_total
from customer
join invoice on customer.customer_id = invoice.customer_id
group by  customer.customer_id
order by invoice_total desc
limit 1;
```

Customer_id	first_name	last_name	invoice_total
5	FrantiÅ¡ek	WichterlovÃ¡	144.54000000000002

- 6) 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.

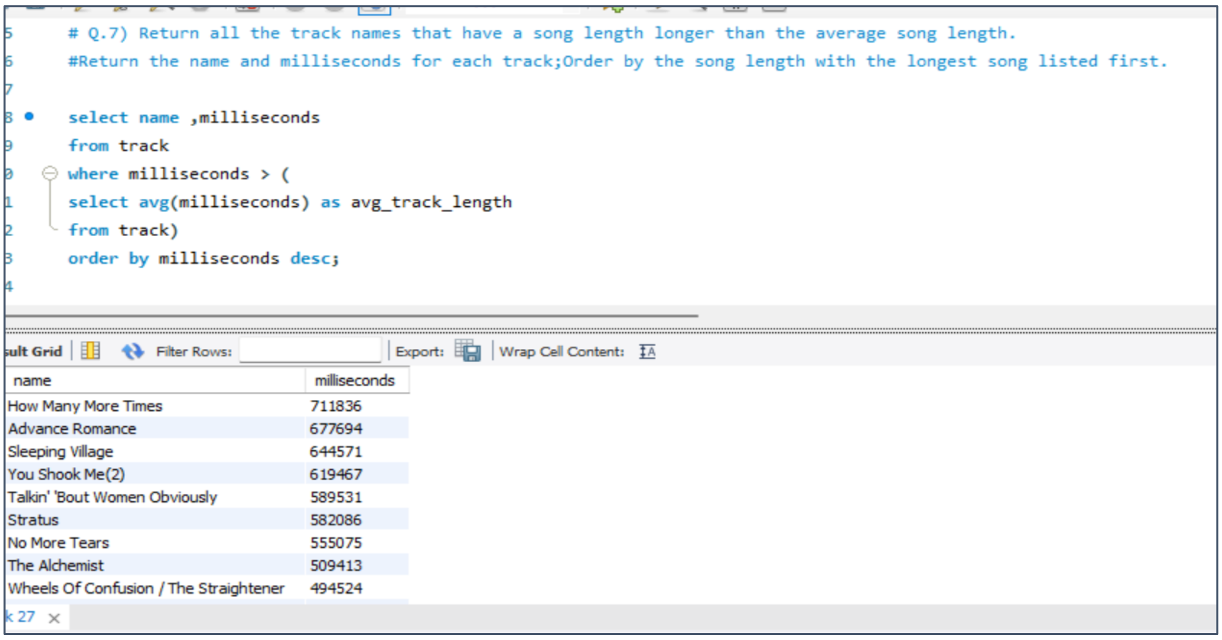


```
# Q.6) Write a query to return the email.first name,last name and genre of all Rock music listeners.
# Return your list ordered alphabetically by email starting with A.

select distinct customer.email as Email,customer.first_name as FirstName,customer.last_name as LastName ,genre.name as 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 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	FirstName	LastName	Name
aaronmitchell@yahoo.ca	Aaron	Mitchell	Rock
alero@uol.com.br	Alexandre	Rocha	Rock
astrid.gruber@apple.at	Astrid	Gruber	Rock
bjorn.hansen@yahoo.no	BjÃ¸rn	Hansen	Rock
camille.bernard@yahoo.fr	Camille	Bernard	Rock
daan_peeters@apple.be	Daan	Peeters	Rock
diego.gutierrez@yahoo.ar	Diego	GutiÃ©rrez	Rock
dmiller@comcast.com	Dan	Miller	Rock
dominiquelefebvre@gmail.com	Dominique	Lefebvre	Rock

- 7) 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 song listed first.

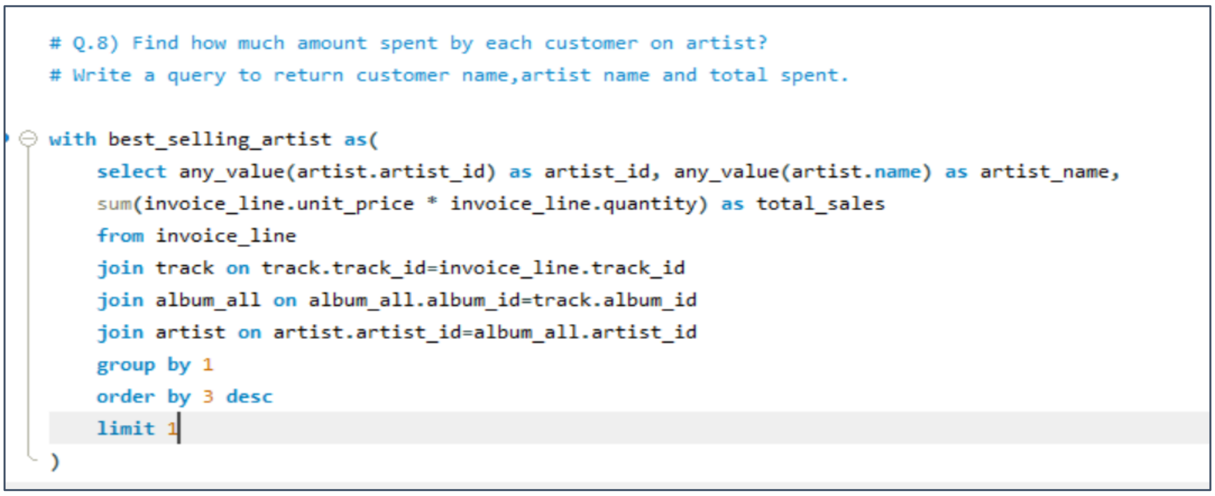


```
# Q.7) Return all the track names that have a song length longer than the average song length.
#Return the name and milliseconds for each track;Order by the song length with the longest song listed first.

select name ,milliseconds
from track
where milliseconds > (
select avg(milliseconds) as avg_track_length
from track)
order by milliseconds desc;
```

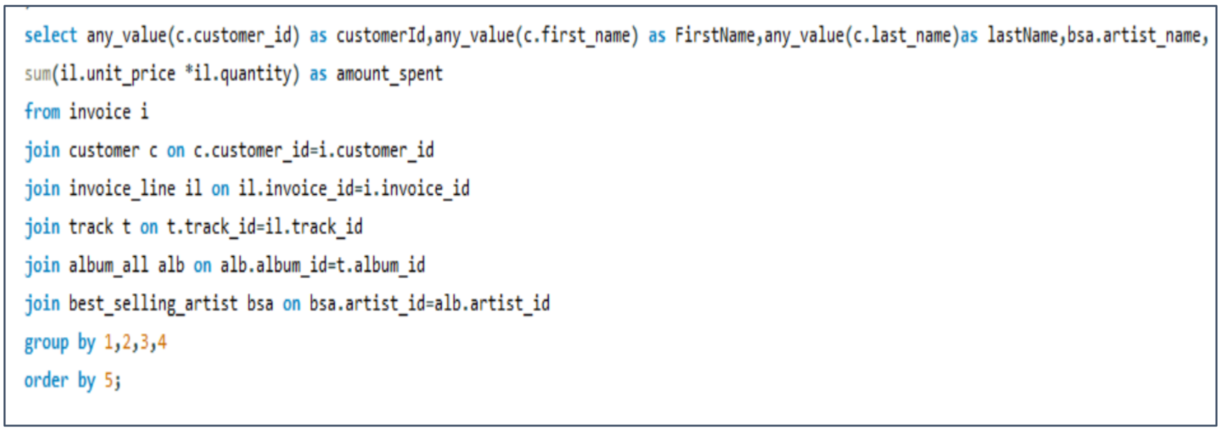
name	milliseconds
How Many More Times	711836
Advance Romance	677694
Sleeping Village	644571
You Shook Me(2)	619467
Talkin' 'Bout Women Obviously	589531
Stratus	582086
No More Tears	555075
The Alchemist	509413
Wheels Of Confusion / The Straightener	494524

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

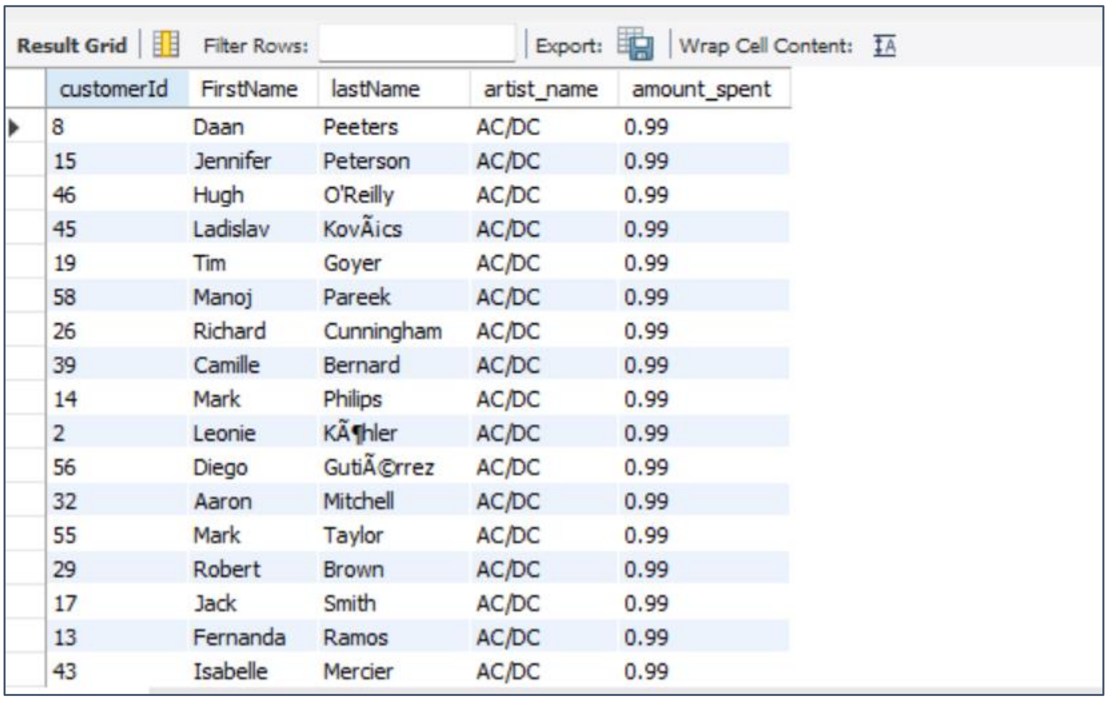


```
# Q.8) Find how much amount spent by each customer on artist?
# Write a query to return customer name,artist name and total spent.

with best_selling_artist as(
    select any_value(artist.artist_id) as artist_id, any_value(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_all on album_all.album_id=track.album_id
    join artist on artist.artist_id=album_all.artist_id
    group by 1
    order by 3 desc
    limit 1
)
```

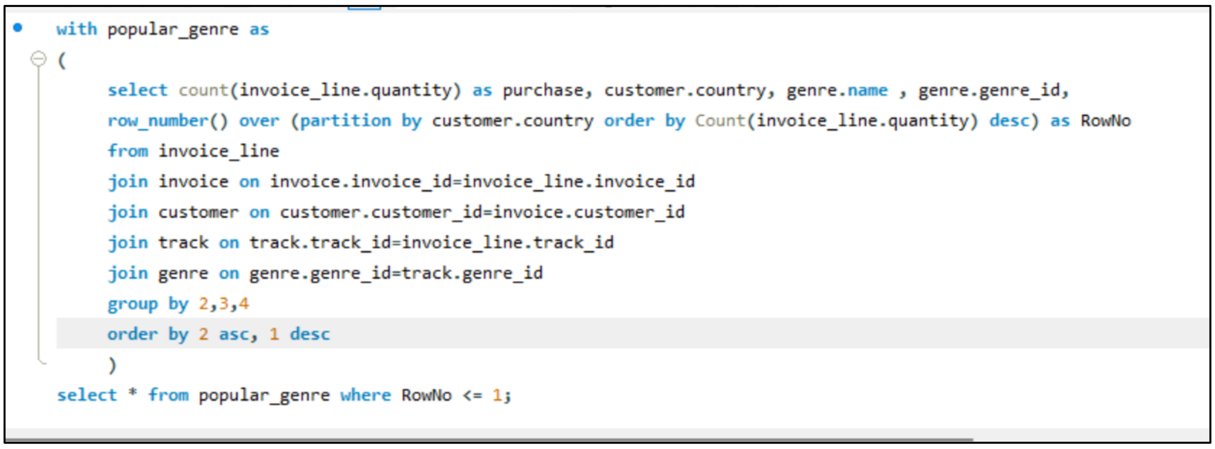


```
select any_value(c.customer_id) as customerId,any_value(c.first_name) as FirstName,any_value(c.last_name)as lastName,bsa.artist_name,
sum(il.unit_price *il.quantity) as amount_spent
from invoice i
join customer c on c.customer_id=i.customer_id
join invoice_line il on il.invoice_id=i.invoice_id
join track t on t.track_id=il.track_id
join album_all 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;
```

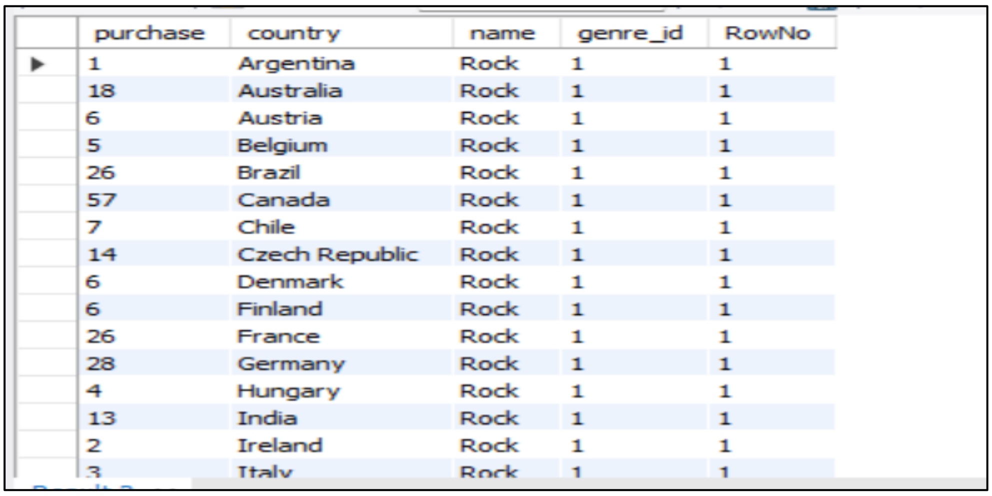


customerId	FirstName	lastName	artist_name	amount_spent
8	Daan	Peeters	AC/DC	0.99
15	Jennifer	Peterson	AC/DC	0.99
46	Hugh	O'Reilly	AC/DC	0.99
45	Ladislav	KovÃ¡cs	AC/DC	0.99
19	Tim	Goyer	AC/DC	0.99
58	Manoj	Pareek	AC/DC	0.99
26	Richard	Cunningham	AC/DC	0.99
39	Camille	Bernard	AC/DC	0.99
14	Mark	Philips	AC/DC	0.99
2	Leonie	KÃ¶hler	AC/DC	0.99
56	Diego	GutiÃ©rrez	AC/DC	0.99
32	Aaron	Mitchell	AC/DC	0.99
55	Mark	Taylor	AC/DC	0.99
29	Robert	Brown	AC/DC	0.99
17	Jack	Smith	AC/DC	0.99
13	Fernanda	Ramos	AC/DC	0.99
43	Isabelle	Mercier	AC/DC	0.99

9) 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 amounts 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.

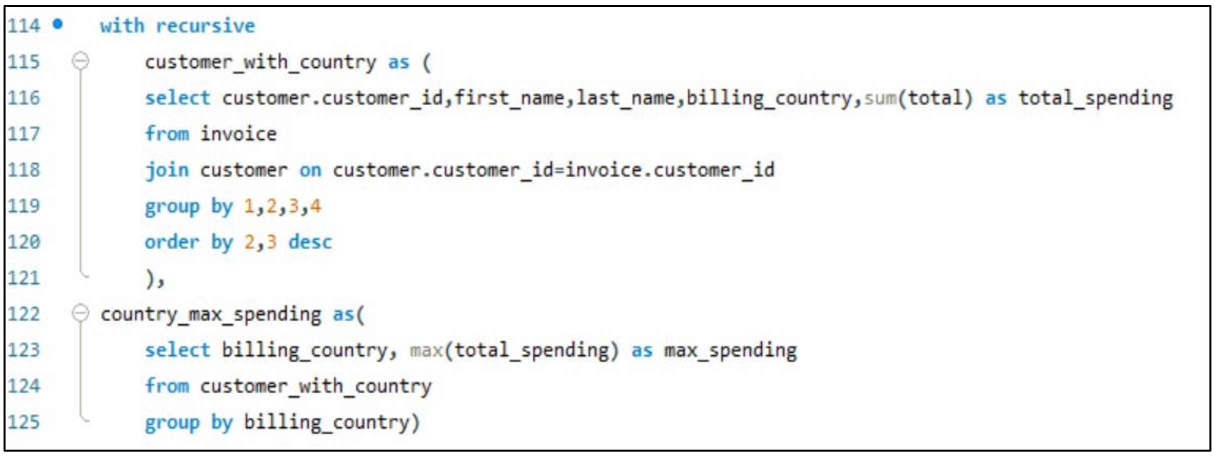


```
with popular_genre as
(
    select count(invoice_line.quantity) as purchase, customer.country, genre.name , genre.genre_id,
    row_number() over (partition by customer.country order by Count(invoice_line.quantity) desc) as RowNo
    from invoice_line
    join invoice on invoice.invoice_id=invoice_line.invoice_id
    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;
```

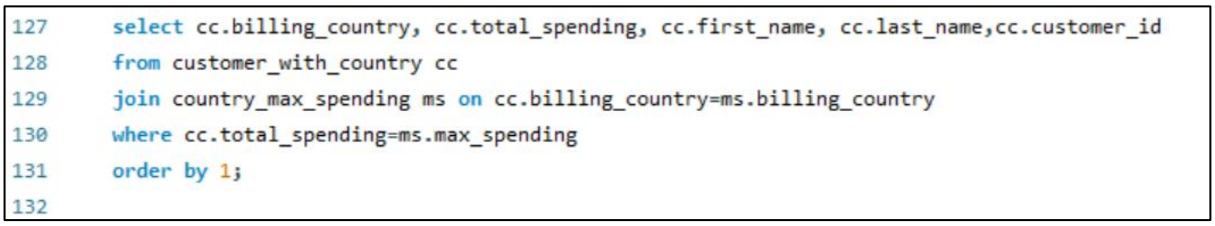


purchase	country	name	genre_id	RowNo
1	Argentina	Rock	1	1
18	Australia	Rock	1	1
6	Austria	Rock	1	1
5	Belgium	Rock	1	1
26	Brazil	Rock	1	1
57	Canada	Rock	1	1
7	Chile	Rock	1	1
14	Czech Republic	Rock	1	1
6	Denmark	Rock	1	1
6	Finland	Rock	1	1
26	France	Rock	1	1
28	Germany	Rock	1	1
4	Hungary	Rock	1	1
13	India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

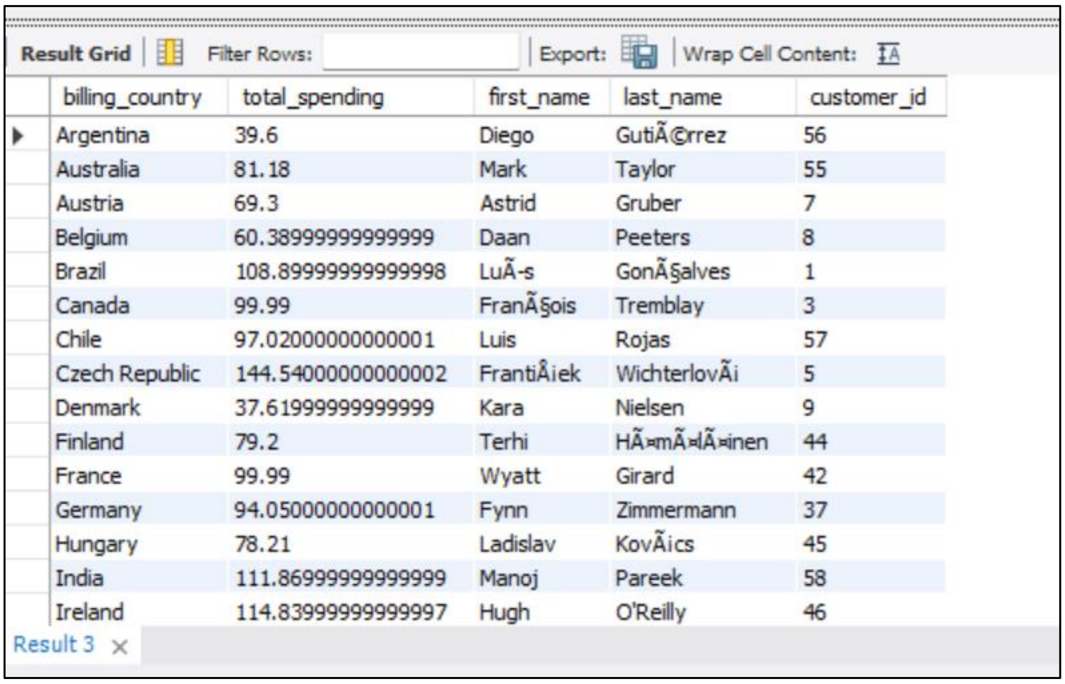
- 10) Write a query that determines the customer that has spent the most on music for each country. Write a query that returns the country along with the top customer and how much they spent. For countries where the top amount spent is shared, provide all customers who spent this amount.



```
with recursive
    customer_with_country as (
    select customer.customer_id,first_name,last_name,billing_country,sum(total) as total_spending
    from invoice
    join customer on customer.customer_id=invoice.customer_id
    group by 1,2,3,4
    order by 2,3 desc
    ),
country_max_spending as(
    select billing_country, max(total_spending) as max_spending
    from customer_with_country
    group by billing_country)
```



```
    select cc.billing_country, cc.total_spending, cc.first_name, cc.last_name,cc.customer_id
    from customer_with_country cc
    join country_max_spending ms on cc.billing_country=ms.billing_country
    where cc.total_spending=ms.max_spending
    order by 1;
```



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

billing_country	total_spending	first_name	last_name	customer_id
Argentina	39.6	Diego	GutiÃ©rrez	56
Australia	81.18	Mark	Taylor	55
Austria	69.3	Astrid	Gruber	7
Belgium	60.38999999999999	Daan	Peeters	8
Brazil	108.89999999999998	LuÃ-s	GonÃ§alves	1
Canada	99.99	FranÃ§ois	Tremblay	3
Chile	97.02000000000001	Luis	Rojas	57
Czech Republic	144.54000000000002	FrantiÅiek	WichterlovÃi	5
Denmark	37.61999999999999	Kara	Nielsen	9
Finland	79.2	Terhi	HÃ¤mÃ¤lÃ¤inen	44
France	99.99	Wyatt	Girard	42
Germany	94.05000000000001	Fynn	Zimmermann	37
Hungary	78.21	Ladislav	KovÃics	45
India	111.86999999999999	Manoj	Pareek	58
Ireland	114.83999999999997	Hugh	O'Reilly	46

Result 3 ×