

PROJECT:- 4

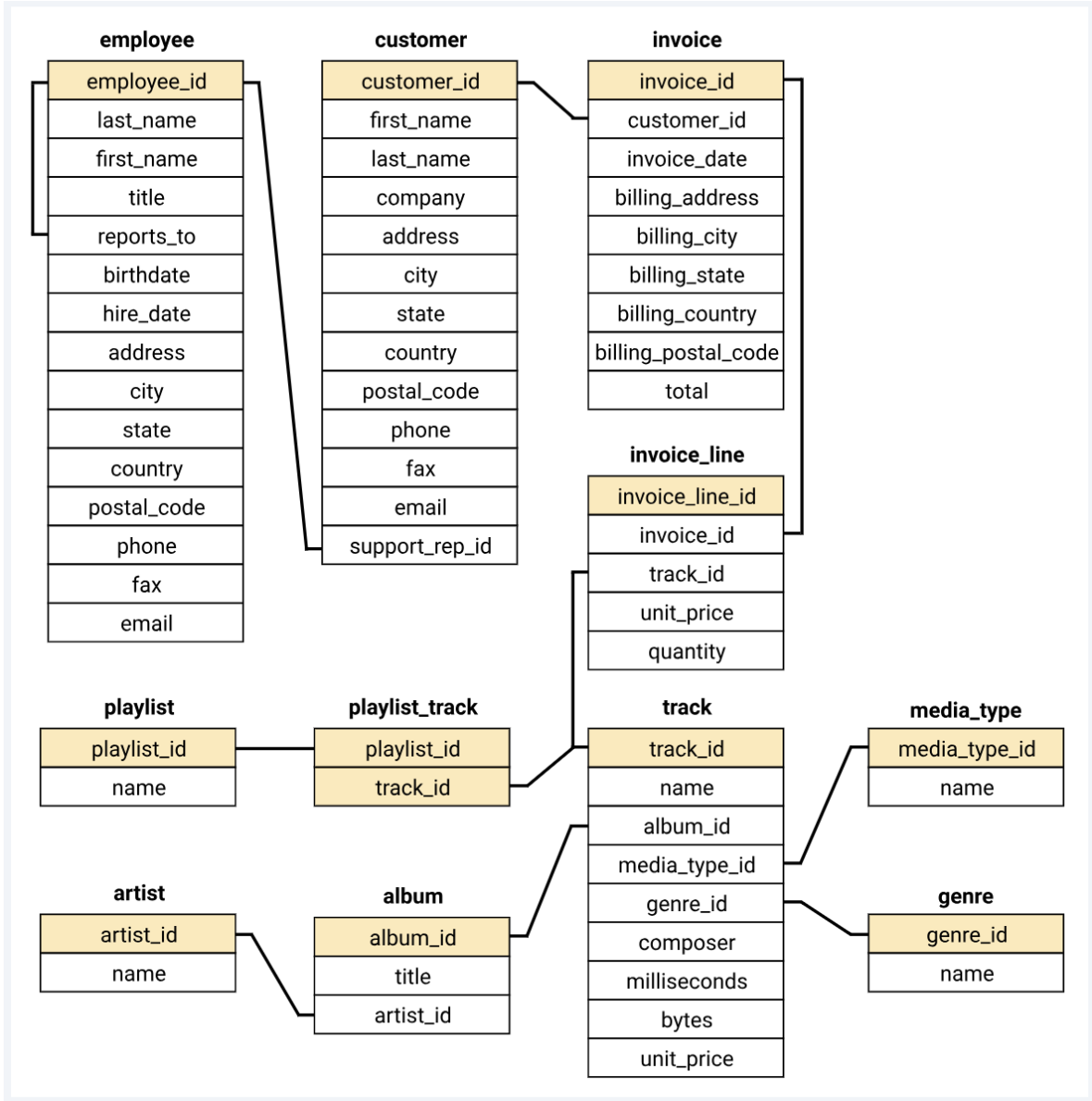
MUSIC STORE DATA ANALYSIS IN SQL

- You are hired as a data analyst at a music store, let's call it "Spotify". You're given access to a database with information about customers, singers, and more.
- As you explore connections between the data, your manager shares exciting news: the company plans to launch a new product! They need insights from the data to guide their decision-making.
- Your task is to analyse the database and provide answers to your manager's problem statements

Problem Statements

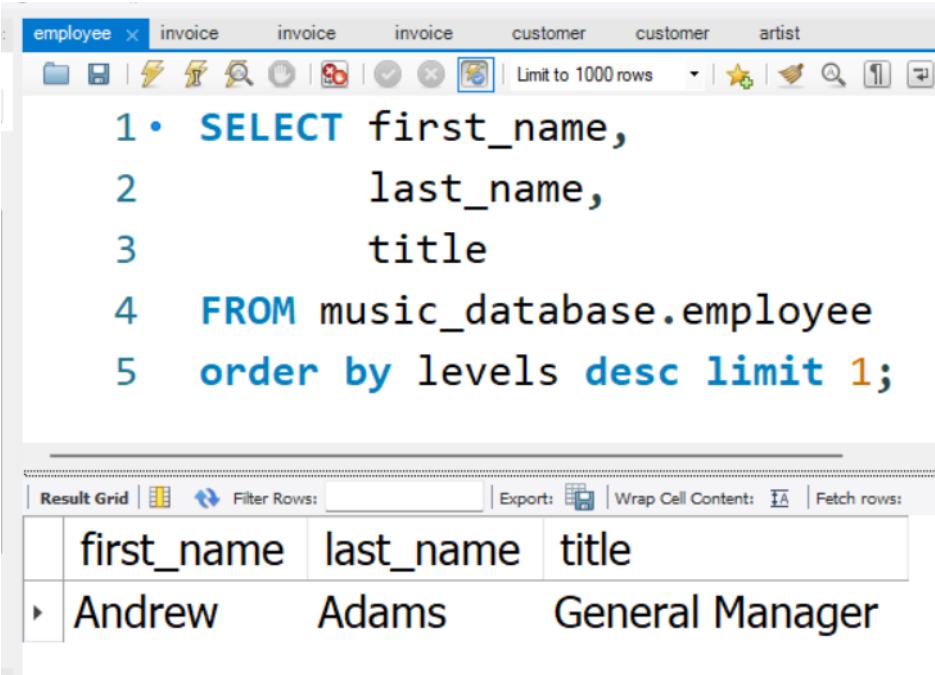
1. We want to assign an senior employee to lead that project. Who is the senior most employee based on job title?
2. Which county has most number of Invoices?
3. What is value of top 3 invoices ?
4. We would like to throw a promotional Music Festival in the city we made the most money. Which city has the best customers?
5. Who is the best customer? The customer who has spent the most money will be declared the best customer.
6. Details of customers who listens Rock music.
7. Let's invite the artists who have written the most rock music in our dataset.
8. We want to find out the most popular music Genre for each country.

Database Schema:



Problem Statements:

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



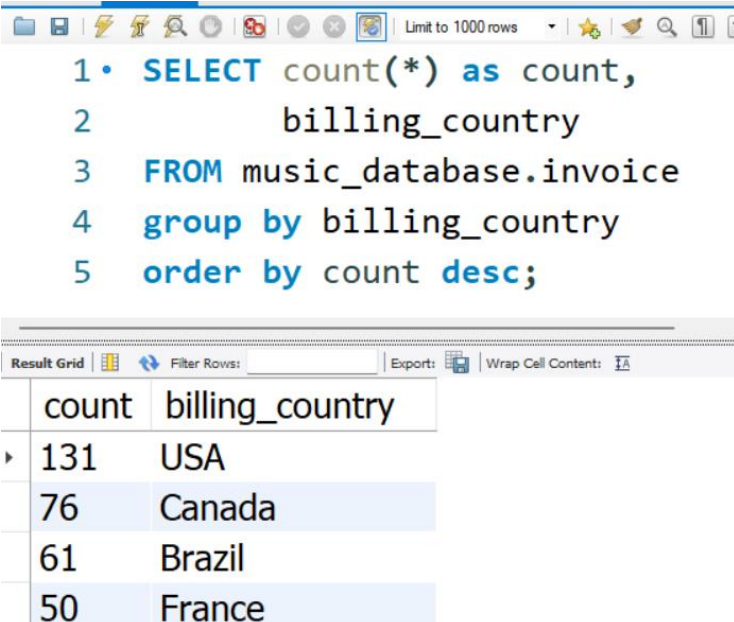
The screenshot shows a database query editor with a toolbar at the top containing icons for file operations, execution, and search. The query is as follows:

```
1 • SELECT first_name,  
2         last_name,  
3         title  
4 FROM music_database.employee  
5 ORDER BY levels DESC LIMIT 1;
```

Below the query editor is a 'Result Grid' with a toolbar for filtering, exporting, and wrapping text. The result is displayed in a table with three columns: first_name, last_name, and title.

first_name	last_name	title
Andrew	Adams	General Manager

2. Which countries have the most Invoices?



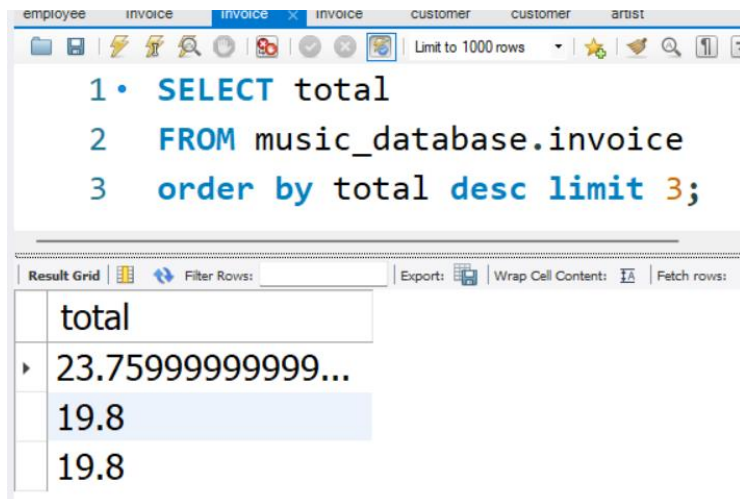
The screenshot shows a database query editor with a toolbar at the top. The query is as follows:

```
1 • SELECT count(*) as count,  
2         billing_country  
3 FROM music_database.invoice  
4 GROUP BY billing_country  
5 ORDER BY count DESC;
```

Below the query editor is a 'Result Grid' with a toolbar for filtering, exporting, and wrapping text. The result is displayed in a table with two columns: count and billing_country.

count	billing_country
131	USA
76	Canada
61	Brazil
50	France

3. What are top 3 values of total invoice?



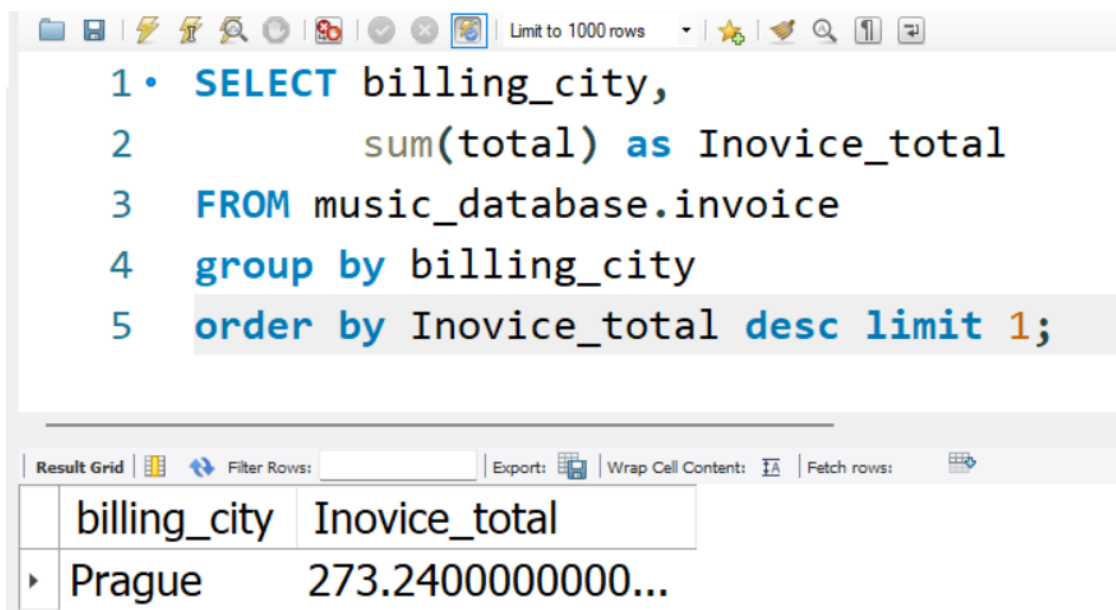
The screenshot shows a database query editor with a toolbar at the top containing icons for file operations, search, and execution. Below the toolbar, a SQL query is entered in a text area. The query is as follows:

```
1 • SELECT total
2 FROM music_database.invoice
3 order by total desc limit 3;
```

Below the query editor, there is a 'Result Grid' section. It includes a 'Filter Rows' input field, an 'Export' button, a 'Wrap Cell Content' checkbox, and a 'Fetch rows' button. The result grid displays the following data:

total
23.759999999999...
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.



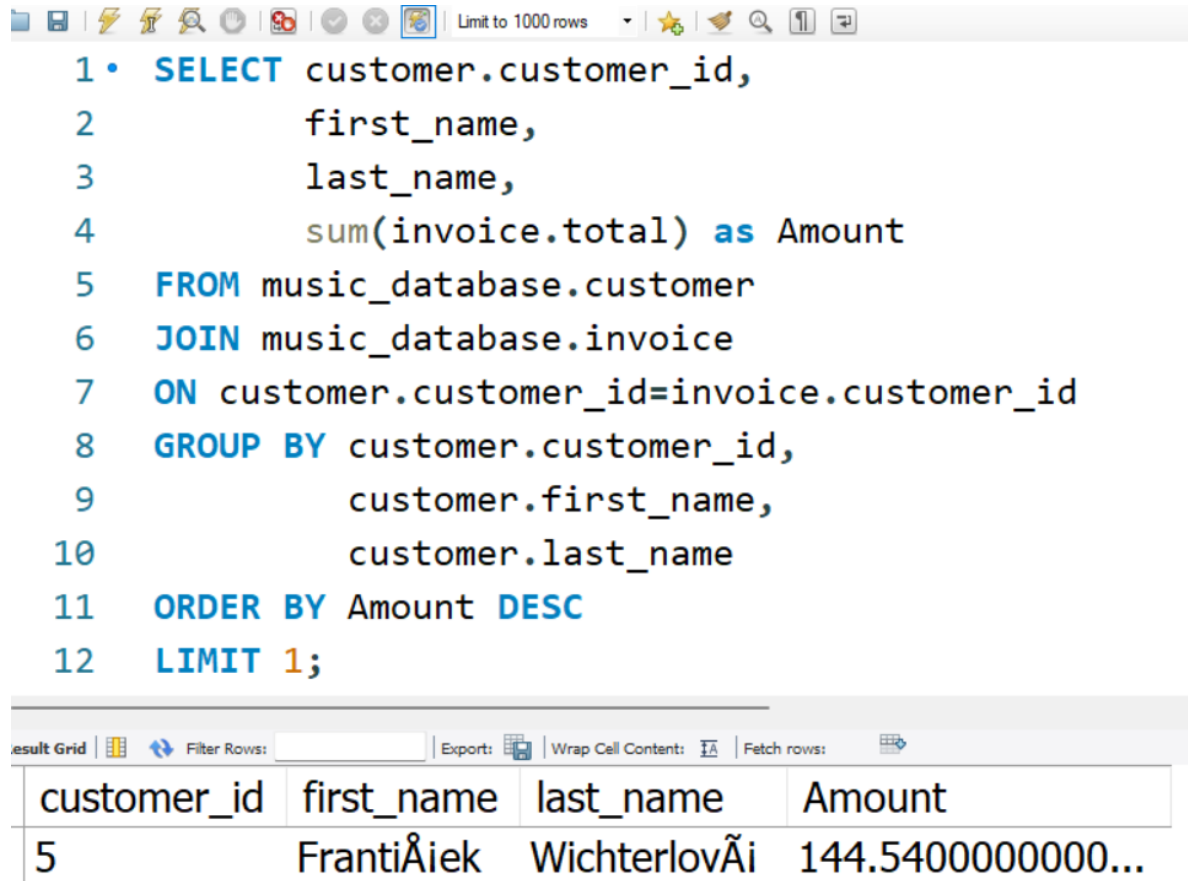
The screenshot shows a database query editor with a toolbar at the top containing icons for file operations, search, and execution. Below the toolbar, a SQL query is entered in a text area. The query is as follows:

```
1 • SELECT billing_city,
2         sum(total) as Invoice_total
3 FROM music_database.invoice
4 group by billing_city
5 order by Invoice_total desc limit 1;
```

Below the query editor, there is a 'Result Grid' section. It includes a 'Filter Rows' input field, an 'Export' button, a 'Wrap Cell Content' checkbox, and a 'Fetch rows' button. The result grid displays the following data:

billing_city	Invoice_total
Prague	273.240000000000...

5. Who is the best customer? The customer who has spent the most money will be declared the best customer.

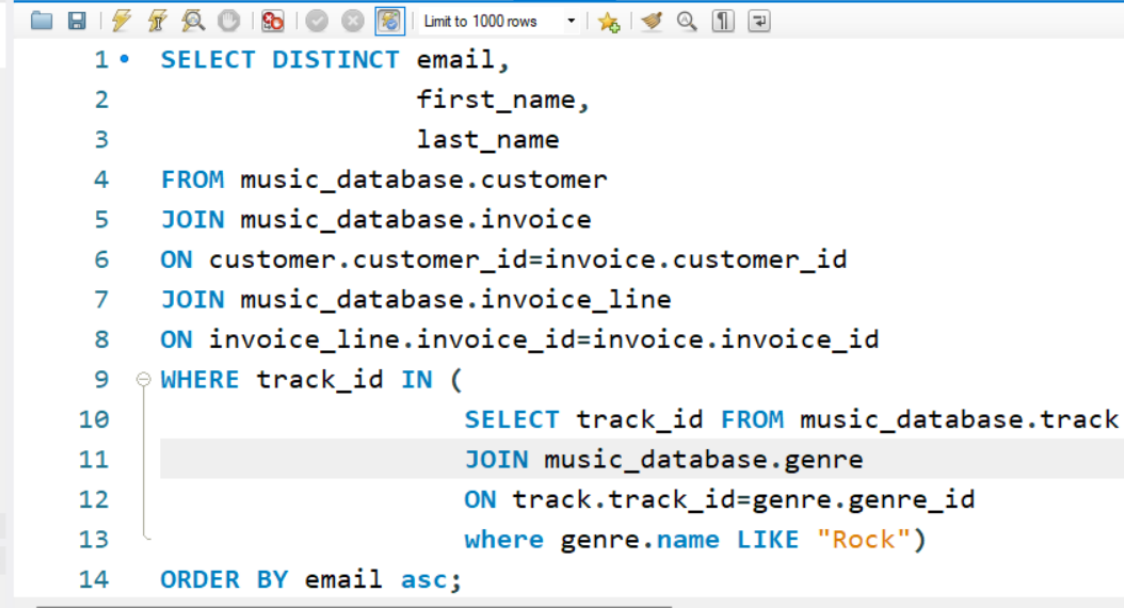


The image shows a SQL query editor window with a toolbar at the top. The query is a SQL statement to find the customer with the highest total spending. Below the query is a results grid showing the output of the query. The results grid has a toolbar with options like 'Filter Rows', 'Export', 'Wrap Cell Content', and 'Fetch rows'. The results are displayed in a table with four columns: customer_id, first_name, last_name, and Amount. The first row shows customer_id 5, first_name František, last_name Wichterlov, and Amount 144.540000000000...

```
1 • SELECT customer.customer_id,  
2         first_name,  
3         last_name,  
4         sum(invoice.total) as Amount  
5 FROM music_database.customer  
6 JOIN music_database.invoice  
7 ON customer.customer_id=invoice.customer_id  
8 GROUP BY customer.customer_id,  
9         customer.first_name,  
10        customer.last_name  
11 ORDER BY Amount DESC  
12 LIMIT 1;
```

customer_id	first_name	last_name	Amount
5	František	Wichterlov	144.540000000000...

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.



```

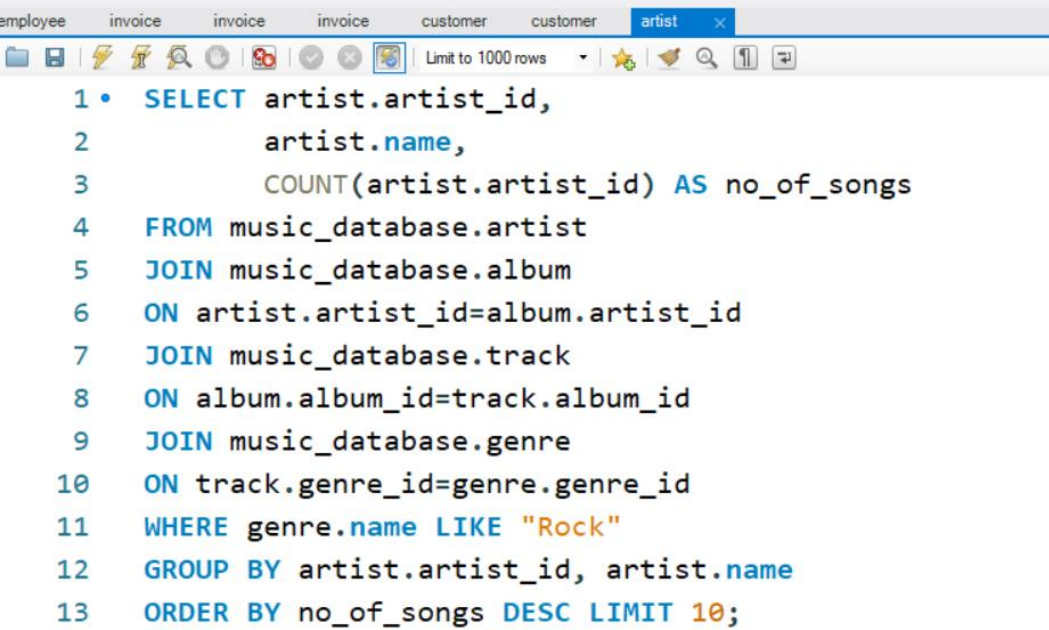
1 • SELECT DISTINCT email,
2     first_name,
3     last_name
4 FROM music_database.customer
5 JOIN music_database.invoice
6 ON customer.customer_id=invoice.customer_id
7 JOIN music_database.invoice_line
8 ON invoice_line.invoice_id=invoice.invoice_id
9 WHERE track_id IN (
10     SELECT track_id FROM music_database.track
11     JOIN music_database.genre
12     ON track.track_id=genre.genre_id
13     where genre.name LIKE "Rock")
14 ORDER BY email asc;

```

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

	email	first_name	last_name
▶	aaronmitchell@y...	Aaron	Mitchell
	johavanderberg...	Johannes	Van der Berg

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.



```

1 • SELECT artist.artist_id,
2     artist.name,
3     COUNT(artist.artist_id) AS no_of_songs
4 FROM music_database.artist
5 JOIN music_database.album
6 ON artist.artist_id=album.artist_id
7 JOIN music_database.track
8 ON album.album_id=track.album_id
9 JOIN music_database.genre
10 ON track.genre_id=genre.genre_id
11 WHERE genre.name LIKE "Rock"
12 GROUP BY artist.artist_id, artist.name
13 ORDER BY no_of_songs DESC LIMIT 10;

```

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

	artist_id	name	no_of_songs
▶	1	AC/DC	18
	3	Aerosmith	15

8. 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

```
1 • WITH popular_genre as
2   (SELECT invoice.billing_country,
3          SUM(invoice_line.quantity) quantity,
4          genre.name as genre,
5
6   ROW_NUMBER()
7   over(PARTITION BY invoice.billing_country
8        ORDER BY SUM(invoice_line.quantity) DESC) AS row_num
9
10  FROM music_database.customer
11      JOIN invoice
12      ON customer.customer_id=invoice.customer_id
13      JOIN invoice_line
14      ON invoice.invoice_id=invoice_line.invoice_id
15      JOIN track
16      ON invoice_line.track_id=track.track_id
17      JOIN genre
18      ON genre.genre_id=track.genre_id
19  GROUP BY 1,3
20  ORDER BY 2 DESC)
21  SELECT * FROM popular_genre
```

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

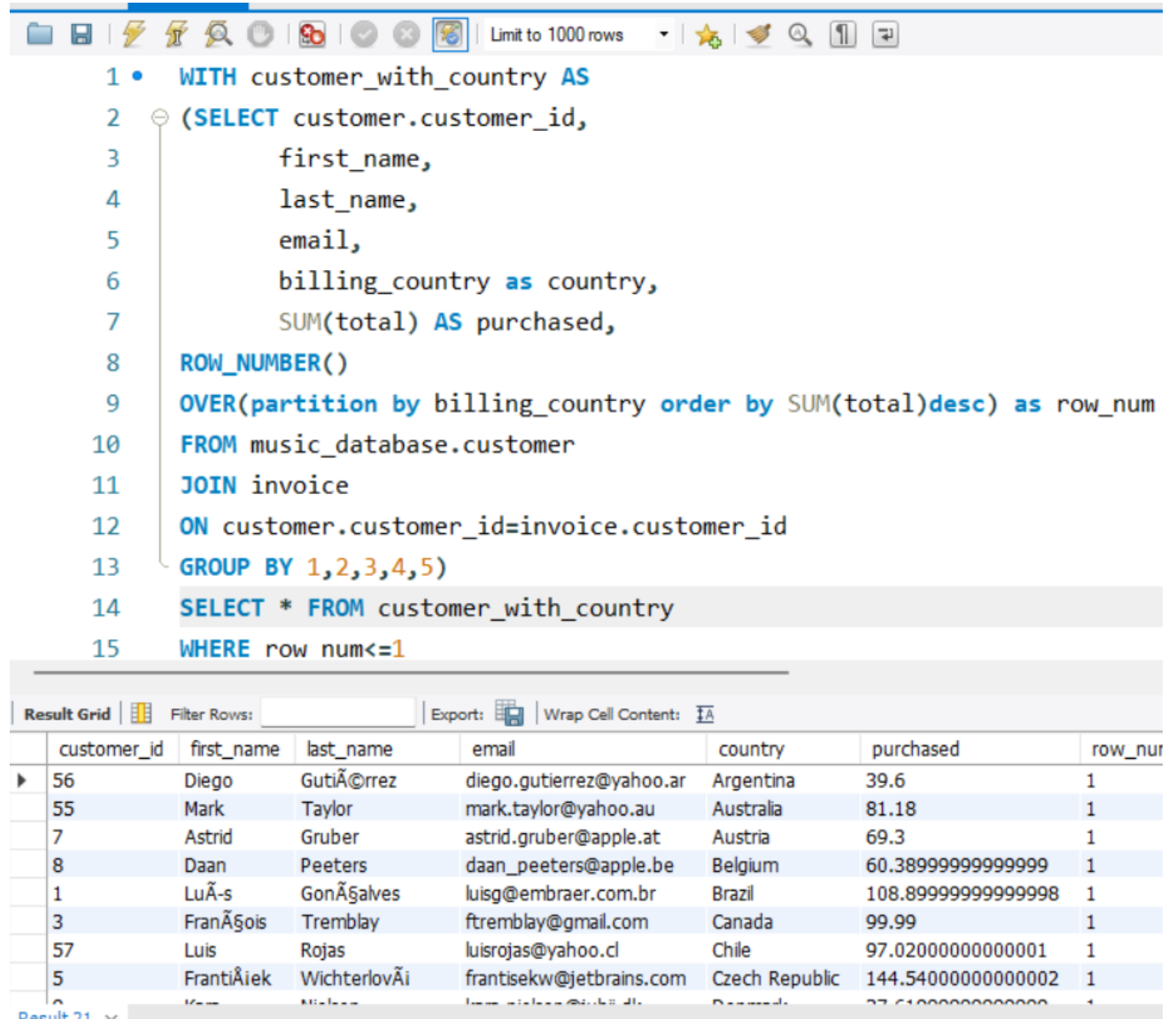
	billing_country	quantity	genre	row_num
▶	USA	70	Rock	1
	Canada	57	Rock	1
	United Kingdom	47	Rock	1
	Germany	28	Rock	1

Result 36 x

9. Determine which customer has spent the most on music for each country.

OR

Write a query that returns the country along with the top customer and how



The screenshot shows a SQL IDE interface. The top toolbar includes icons for file operations, execution, and a 'Limit to 1000 rows' dropdown. The SQL editor contains the following query:

```
1 • WITH customer_with_country AS
2   (SELECT customer.customer_id,
3          first_name,
4          last_name,
5          email,
6          billing_country as country,
7          SUM(total) AS purchased,
8          ROW_NUMBER()
9          OVER(partition by billing_country order by SUM(total)desc) as row_num
10  FROM music_database.customer
11  JOIN invoice
12  ON customer.customer_id=invoice.customer_id
13  GROUP BY 1,2,3,4,5)
14  SELECT * FROM customer_with_country
15  WHERE row num<=1
```

Below the query editor is the 'Result Grid' section, which includes a 'Filter Rows' input, 'Export' and 'Wrap Cell Content' buttons, and a table of results. The table has 8 columns: customer_id, first_name, last_name, email, country, purchased, and row_num. The results show the top customer for each country, with the top customer for each country having a row_num of 1.

customer_id	first_name	last_name	email	country	purchased	row_num
56	Diego	Gutiérrez	diego.gutierrez@yahoo.ar	Argentina	39.6	1
55	Mark	Taylor	mark.taylor@yahoo.au	Australia	81.18	1
7	Astrid	Gruber	astrid.gruber@apple.at	Austria	69.3	1
8	Daan	Peeters	daan_peeters@apple.be	Belgium	60.38999999999999	1
1	Luís	Gonçalves	luisg@embraer.com.br	Brazil	108.89999999999998	1
3	François	Tremblay	ftremblay@gmail.com	Canada	99.99	1
57	Luis	Rojas	luisrojas@yahoo.cl	Chile	97.02000000000001	1
5	František	Wichterlov	frantisekw@jetbrains.com	Czech Republic	144.54000000000002	1
6	Kari	Niemi	kari.niemi@iki.fi	Finland	37.61000000000000	1