DIGITAL MUSIC STORE ANALYSIS

Bullet Points to analyse the Music store database:

- **Senior most employee based on job title.**
- **Countries with most number of invoices.**
- **Maximum, Minimum and Average amount of invoice.**
- **City that has the highest sum of invoice totals.**
- **The customer who has spent the most money.**
- **Email, first name, last name, & Genre of all Rock Music listeners.**
- **Rock bands who have written the most rock music tracks.**
- **Highest amount spent by Top 10 customer on which artists?**
- **Most popular music Genre for each country.**
- **Customer that has spent the most on music from each country.**

Data Analysed by performing SQL queries using Joins, Window functions and CTE.

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

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

employee_id [PK] character varying (50)	first_name character (50)	last_name character (50)	title character varying (50)	levels character varying (10)
9	Mohan	Madan	Senior General Manager	L7

The senior most employee is Mohan Madan, who is Senior General Manager.

2. Which countries have the most Invoices (TOP 10)?

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

Result**:** Below-mentioned countries have the most invoices, with USA, Canada and Brazil being the top 3 countries.

	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

3. What are the values of maximum amount of invoice, minimum amount of invoice and average amount of invoice?

```
SELECT
      MAX(total),
      MIN(total),
      AVG(total)
FROM invoice;
```

Result: Below are the values of maximum, minimum and average amount of invoice.

	max real	min real	avg double precision
1	23.76	0.99	7.67008139722122

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

	billing_city character varying (30)	invoicetotal real
1	Prague	273.23996

Prague is the city with best customers with the sum of all invoices from the city more than $273.

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 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_i d
ORDER BY total_spending DESC
LIMIT 1;
```

	customer_id [PK] integer	first_name character (50)	last_name character (50)	country character varying (50)	total_spending real
1	5	R ...	Madhav	Czech Republic	144.54002

Best customer is R Madhav from Czech Republic who spent around $144.5 overall.

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.

```
SELECT DISTINCT email,first_name, last_name
FROM customer
JOIN invoice ON customer.customer_id = invoice.customer_id
JOIN invoiceline ON invoice.invoice_id = invoiceline.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;
```

Result: Below are the details of Rock music listeners. There are 59 listeners of which 15 are shown below.

	email character varying (50)	first_name character (50)	last_name character (50)
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@yachoo.ca	Edward	Francis
11	eduardo@woodstock.com.br	Eduardo	Martins
12	ellie.sullivan@shaw.ca	Ellie	Sullivan
13	emma_jones@hotmail.com	Emma	Jones
14	enrique_munoz@yahoo.es	Enrique	Muñoz
15	fernadaramos4@uol.com.br	Fernanda	Ramos

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.

```
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;
```

Result: Below are the details of Top 10 Artists who have written the most rock music tracks

	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

8. Find how much amount spent by Top 10 customer on best-selling artist? Write a query to return customer name, artist name and total spent.

Steps to Solve:
(i) Find artist who has earned the most according to the Invoice lines.
(ii) Find which customer spent the most on that artist.
.

```
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;
```

Result: Below are the Top 10 customers, who spend highest on the best-selling customer, with the customer_id and the amount_spend.

	customer_id integer	first_name character (50)	last_name character (50)	artist_name character varying (120)	amount_spent 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	41	Marc	Dubois	Queen	11.88
6	53	Phil	Hughes	Queen	11.88
7	33	Ellie	Sullivan	Queen	10.89
8	47	Lucas	Mancini	Queen	10.89
9	20	Dan	Miller	Queen	3.96
10	5	R	Madhav	Queen	3.96

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 number 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: There are two parts in question
(i) Find most popular music genre
(ii) Get data at country level.

```
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
```

Result: Below are the 24 countries with most popular music genre.

	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	211	France	Rock	1	1
12	194	Germany	Rock	1	1

13	44	Hungary	Rock	1	1
14	102	India	Rock	1	1
15	72	Ireland	Rock	1	1
16	35	Italy	Rock	1	1
17	33	Netherlands	Rock	1	1
18	40	Norway	Rock	1	1
19	40	Poland	Rock	1	1
20	108	Portugal	Rock	1	1
21	46	Spain	Rock	1	1
22	60	Sweden	Rock	1	1
23	166	United Kingdom	Rock	1	1
24	561	USA	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.

Steps to Solve: Similar to the above question. There are two parts in question:
(i) Find the most spent on music for each country
(ii) Filter the data for respective customers.

```
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
```

Result: Below are the 24 countries and their top spending customers.

	customer_id integer	first_name character (50)	last_name character (50)	billing_country character varying (30)	total_spending real	rowno bigint
1	56	Diego	Gutiérrez	Argentina	39.600002	1
2	55	Mark	Taylor	Australia	81.18	1
3	7	Astrid	Gruber	Austria	69.299995	1
4	8	Daan	Peeters	Belgium	60.39	1
5	1	Luís	Gonçalves	Brazil	108.90001	1
6	3	François	Tremblay	Canada	99.990005	1
7	57	Luis	Rojas	Chile	97.020004	1
8	5	R	Madhav	Czech Republic	144.54002	1
9	9	Kara	Nielsen	Denmark	37.62	1
10	44	Terhi	Hämäläinen	Finland	79.2	1
11	42	Wyatt	Girard	France	99.990005	1
12	37	Fynn	Zimmermann	Germany	94.049995	1

13	45	Ladislav	...	Kovács	...	Hungary	78.21001	1
14	58	Manoj	...	Pareek	...	India	111.869995	1
15	46	Hugh	...	O'Reilly		Ireland	114.84	1
16	47	Lucas	...	Mancini	...	Italy	50.49	1
17	48	Johannes	...	Van der Berg	...	Netherlands	65.34	1
18	4	Bjørn		Hansen	...	Norway	72.27	1
19	49	Stanisław	...	Wójcik	...	Poland	76.229996	1
20	34	João	...	Fernandes	...	Portugal	102.95999	1
21	50	Enrique	...	Muñoz	...	Spain	98.009995	1
22	51	Joakim	...	Johansson	...	Sweden	75.240005	1
23	53	Phil		Hughes	...	United Kingdom	98.01	1
24	17	Jack		Smith		USA	98.01001	1

****