MUSIC STORE DATA ANALYSIS

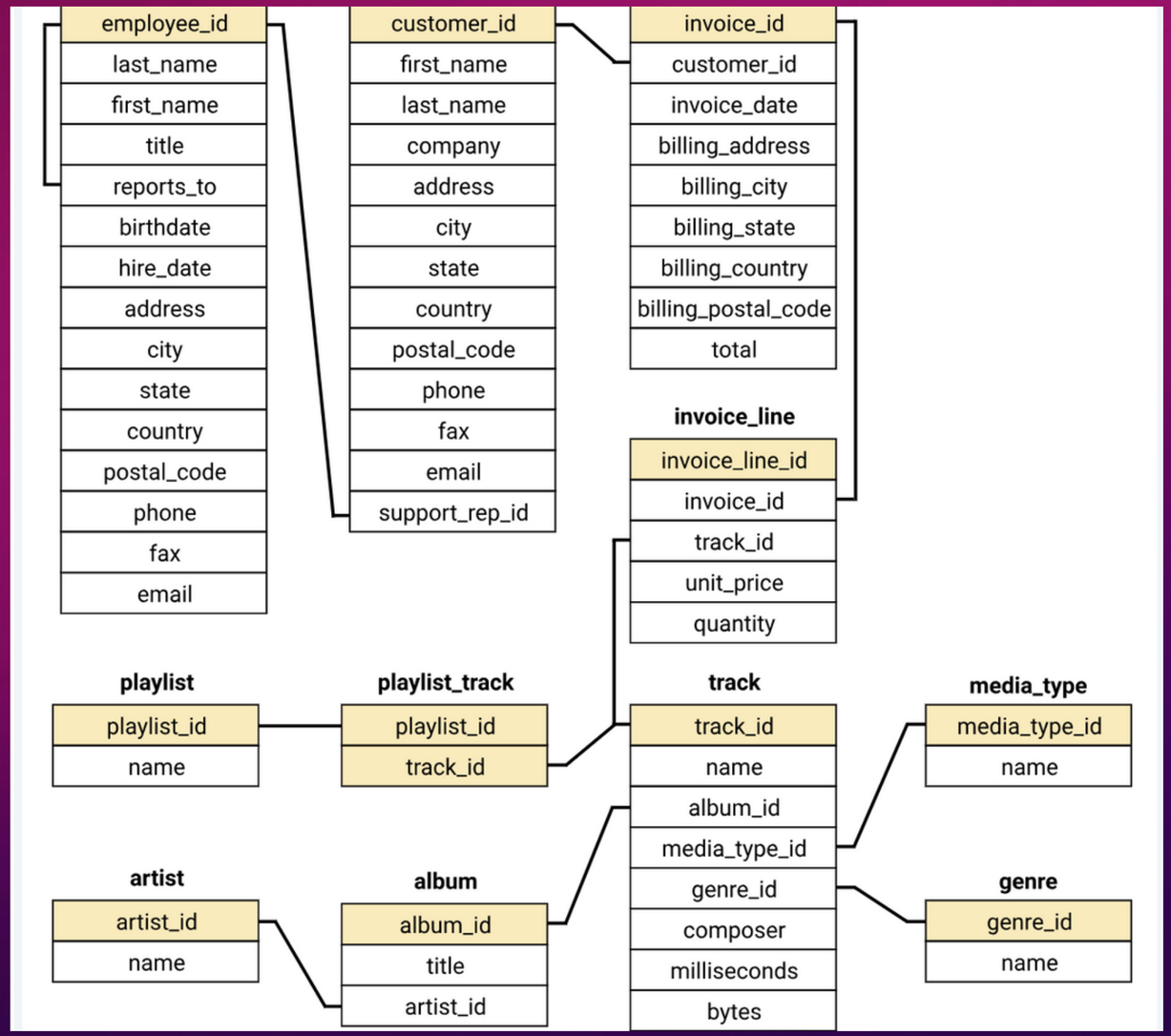
~A SQL Case Study

MySQL™

Problem Statement:

As a data analyst, harness the power of SQL to glean actionable insights from a music store's playlist dataset. Navigate through trends, unearth customer preferences, and illuminate areas for improvement. Tackle business challenges, spotlight top spenders, and distill key insights for crafting impactful offers. Your strategic analysis will drive informed decision-making, shaping the music store's path to sustained growth.

Understanding the Data Set:



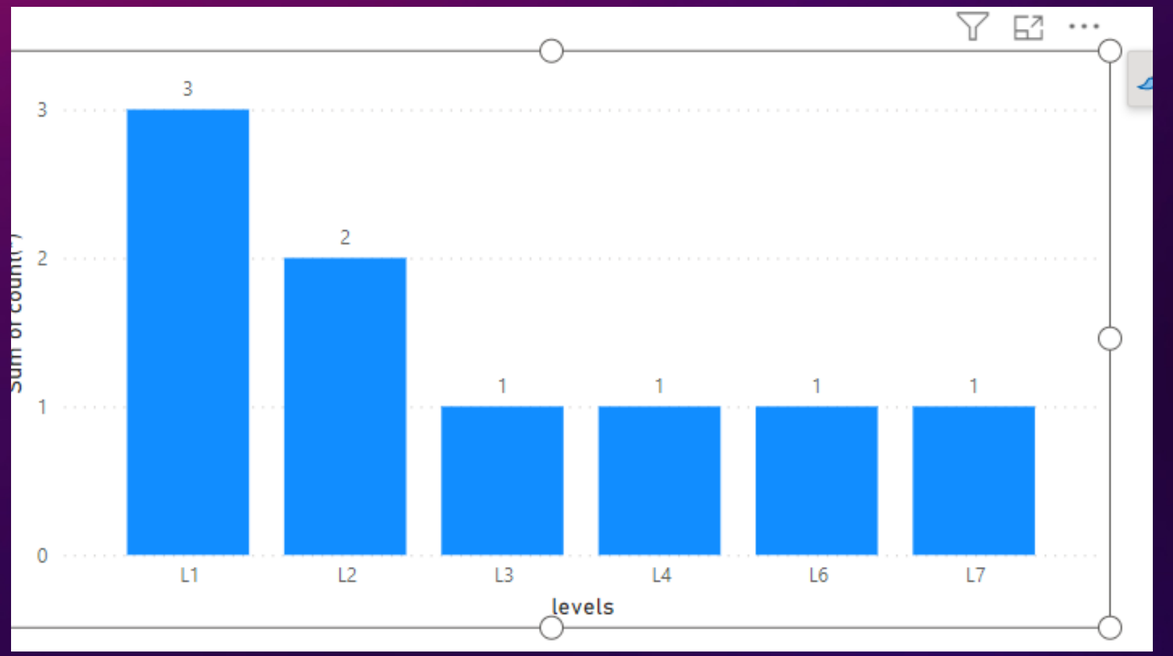
Q1: Who is the senior most employee based on job title?

Query:

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

Output:

	title	last_name	first_name
▶	Senior General Manager	Madan	Mohan



Insights:

1)Clearly,Madan Mohan is the senior most employee.

2)No. of employees are less at senior level as compared to junior level

Q2: Which countries have the most Invoices?

Query:

```
#2
Select billing_country,COUNT(invoice_id) AS No_of_invoices from invoice
GROUP BY billing_country
ORDER BY No_of_invoices DESC;
```

Output:

Result Grid | Filter Rows:

billing_country	No_of_invoices
USA	131
Canada	76
Brazil	61
France	50
Germany	41
Czech Republic	30
Portugal	29
United Kingdom	28
India	21
Ireland	13
Chile	13
Finland	11

Insights:

1)USA has the most invoices.

2)However , it can be clearly seen that Europe played a major role in invoice generation as maximum countries are european.

Q3: What are top 3 values of total invoice?

Query:

```sql
#3
SELECT billing_country,billing_city ,total
FROM invoice
ORDER BY total DESC
limit 3;
```

Output:

Result Grid | Filter Rows: | Export:

billing_country	billing_city	total
France	Bordeaux	23.759999999999998
Canada	Winnipeg	19.8
Czech Republic	Prague	19.8

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

Query:

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

Output:

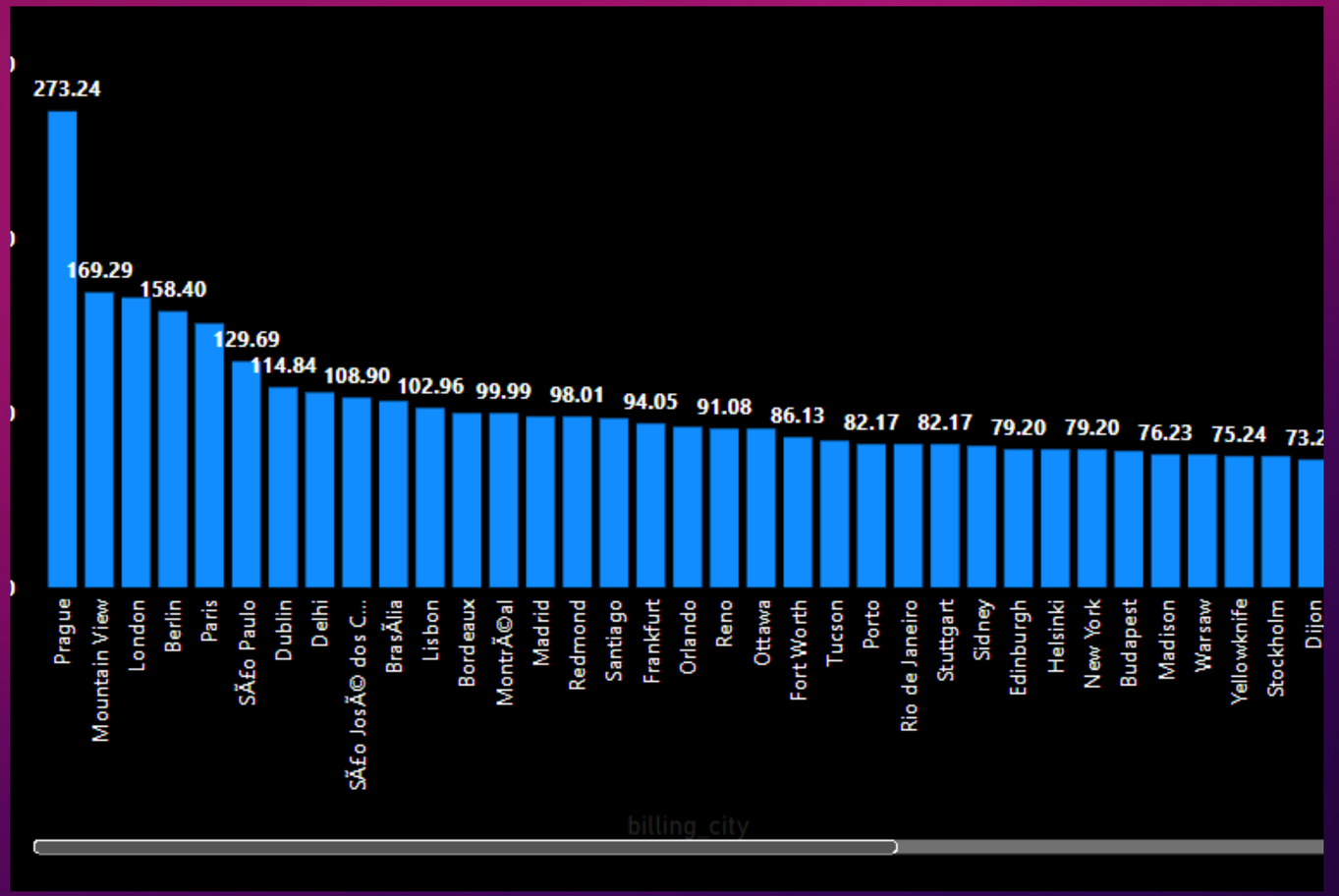
Result Grid | Filter Rows:

billing_city	InvoiceTotal
Prague	273.2400000000007

Insights:

1)Prague has has the highest sum of invoice totals.Hence, is the city with best customers.

2)The other top countries are:Mountain view,London ,Berlin and can also be considered for a promotional festival.

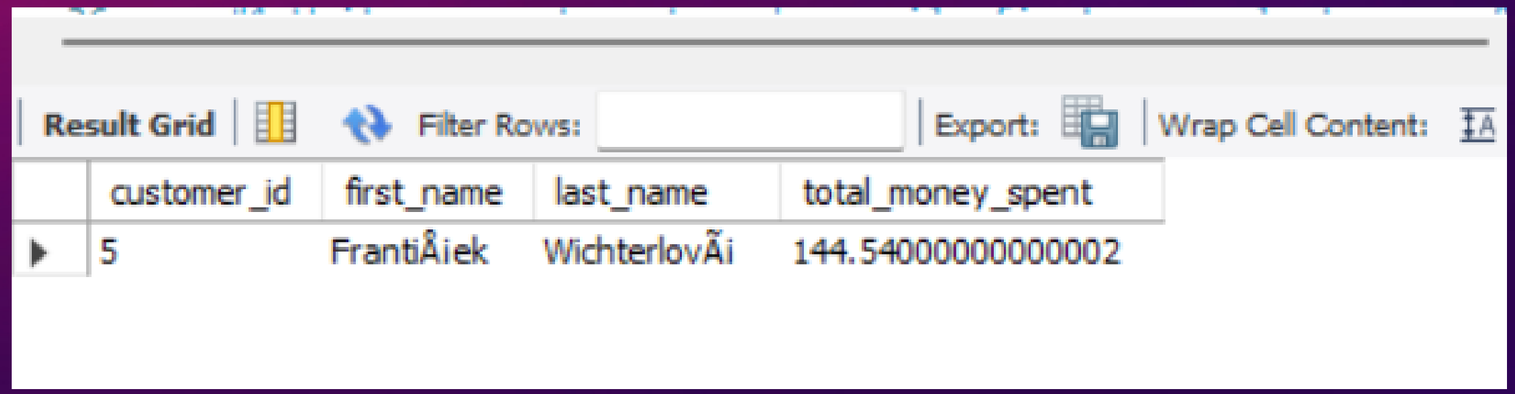


Q 5: Who is the best customer? The customer who has spent the most money will be declared the best customer. Write a query that returns the person who has spent the most money.

Query:

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

Output:



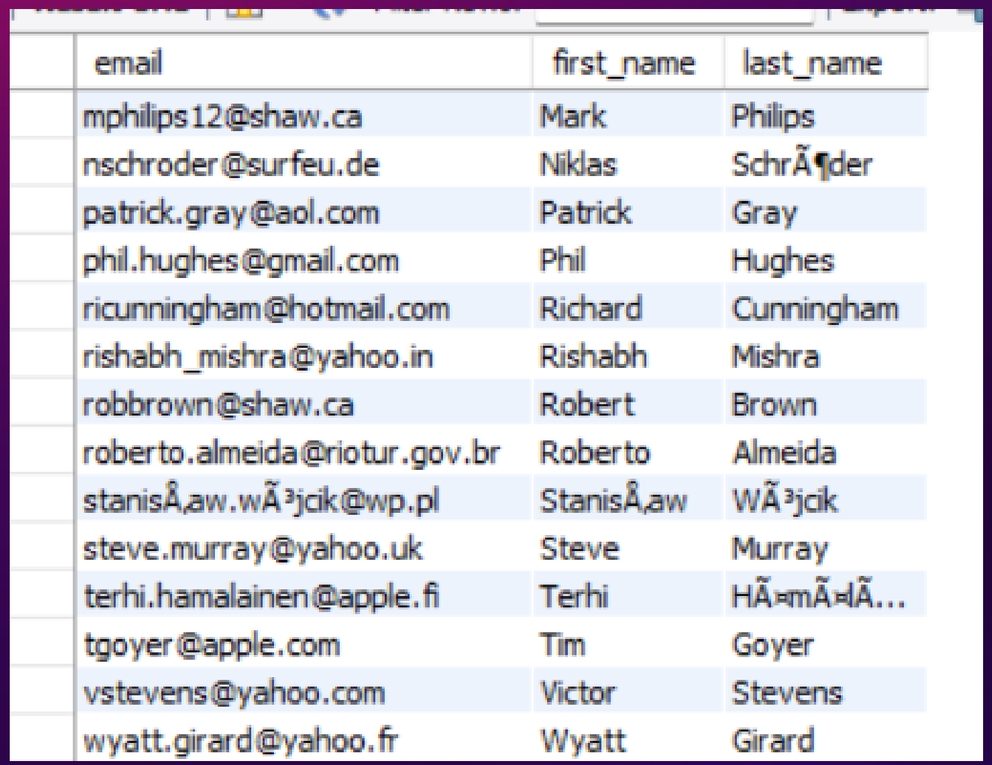
customer_id	first_name	last_name	total_money_spent
5	FrantiÅiek	WichterlovÃi	144.54000000000002

Q 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.

Query:

```
SELECT DISTINCT email,first_name, last_name
FROM customer c
JOIN invoice i ON c.customer_id = i.customer_id
JOIN invoice_line il ON i.invoice_id = il.invoice_id
JOIN track t ON t.track_id = il.track_id
JOIN genre g ON g.genre_id = g.genre_id
WHERE g.name LIKE 'Rock'
ORDER BY email;
```

Output:



email	first_name	last_name
mphilips12@shaw.ca	Mark	Philips
nschroder@surfeu.de	Niklas	SchrÃ¶der
patrick.gray@aol.com	Patrick	Gray
phil.hughes@gmail.com	Phil	Hughes
ricunningham@hotmail.com	Richard	Cunningham
rishabh_mishra@yahoo.in	Rishabh	Mishra
robbrown@shaw.ca	Robert	Brown
roberto.almeida@riotur.gov.br	Roberto	Almeida
stanisÅ‚aw.wÃ³jcik@wp.pl	StanisÅ‚aw	WÃ³jcik
steve.murray@yahoo.uk	Steve	Murray
terhi.hamalainen@apple.fi	Terhi	HÃ¤mÃ¤lÃ…
tgoyer@apple.com	Tim	Goyer
vstevens@yahoo.com	Victor	Stevens
wyatt.girard@yahoo.fr	Wyatt	Girard

Q7: 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

Query:

```
SELECT a.artist_id, a.name,COUNT(a.artist_id) AS number_of_songs
FROM artist a
JOIN album2 al ON al.artist_id = a.artist_id
JOIN track t ON al.album_id = t.album_id
JOIN genre g ON g.genre_id = t.genre_id
WHERE g.name LIKE 'Rock'
GROUP BY a.artist_id
ORDER BY number_of_songs DESC
LIMIT 10;
```

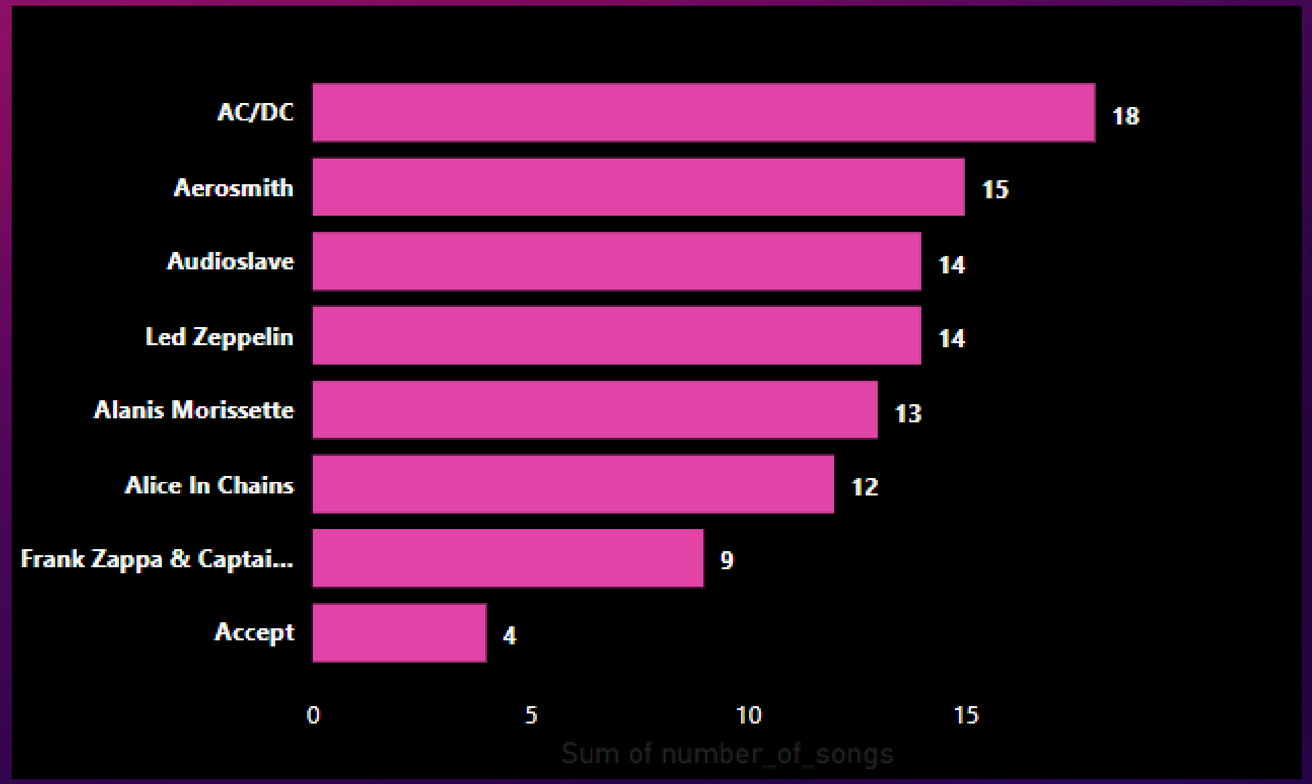
Output:

Result Grid | Filter Rows: | Export:

artist_id	name	number_of_songs
1	AC/DC	18
3	Aerosmith	15
8	Audioslave	14
22	Led Zeppelin	14
4	Alanis Morissette	13
5	Alice In Chains	12
23	Frank Zappa & Captain Beefheart	9
2	Accept	4

Insights:

1)CLearly, AC/DC is the artist who have written the most rock music followed by Aerosmith



Q8: 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.

Query:

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

Output:

name	song_length
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
Book Of Thel	494393
You Oughta Know (Alternate)	491885
Terra	482429

Q9: Find how much amount spent by each customer on artists? Write a query to return customer name, artist name and total spent.

Query:

```
#1
WITH most_earned_artist AS (
    SELECT a.artist_id AS artist_id, a.name AS artist_name, SUM(il.unit_price*il.quantity) AS total_sales
    FROM artist a
    JOIN album2 al ON al.artist_id = a.artist_id
    JOIN track t ON al.album_id = t.album_id
    JOIN invoice_line il ON t.track_id = il.track_id
    GROUP BY 1
    ORDER BY 3 DESC
    LIMIT 1
)

SELECT c.customer_id, c.first_name, c.last_name, m.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 album2 al ON al.album_id = t.album_id
JOIN  most_earned_artist m ON m.artist_id = m.artist_id
GROUP BY 1,2,3,4
ORDER BY 5 DESC;
```

Output:

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

customer_id	first_name	last_name	artist_name	amount_spent
49	StanisÅ‚aw	WÃ³jcik	AC/DC	26.729999999999986
30	Edward	Francis	AC/DC	25.739999999999988
54	Steve	Murray	AC/DC	21.779999999999994
15	Jennifer	Peterson	AC/DC	20.789999999999996
21	Kathy	Chase	AC/DC	20.789999999999996
2	Leonie	KÃ¶hler	AC/DC	19.799999999999997
13	Fernanda	Ramos	AC/DC	18.81
55	Mark	Taylor	AC/DC	18.81
3	FranÃ§ois	Tremblay	AC/DC	18.81
34	JoÃ£o	Fernandes	AC/DC	17.82
24	Frank	Ralston	AC/DC	17.82
1	LuÃ-s	GonÃ§alves	AC/DC	17.82
53	Phil	Hughes	AC/DC	16.830000000000002
52	Emma	Jones	AC/DC	16.830000000000002

Q10: 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.

Query:

```
WITH most_popular_genre AS
(    SELECT COUNT(il.quantity) AS purchases, c.country, g.name, g.genre_id,
    ROW_NUMBER() OVER(PARTITION BY c.country ORDER BY COUNT(il.quantity) DESC) AS rn
    FROM invoice_line il
    JOIN invoice i ON i.invoice_id = il.invoice_id
    JOIN customer c ON c.customer_id = i.customer_id
    JOIN track t ON t.track_id = il.track_id
    JOIN genre g ON g.genre_id = t.genre_id
    GROUP BY 2,3,4
    ORDER BY 2 ASC, 1 DESC  )
SELECT * FROM most_popular_genre WHERE rn <= 1
```

Output:

Result Grid | Filter Rows: | Export:

purchases	country	name	genre_id	rn
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

Insights:

1)Rock is the most popular Genre

Q.11: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

: Query

```sql
WITH most_spent_Customter_with_country AS (
        SELECT c.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 rn
        FROM invoice i
        JOIN customer c ON c.customer_id = i.customer_id
        GROUP BY 1,2,3,4
        ORDER BY 4 ASC,5 DESC)
SELECT * FROM most_spent_Customter_with_country  WHERE rn <= 1
```

Output:

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

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

CONCLUSIONS:

1)Employee Hierarchy Gap:There are fewer employees at the senior level as compared to the junior level

2)Customer Demographics:
The USA dominates in invoice count, but the pivotal role of Europe is evident, with a majority of customers (approximately 70%) stemming from European countries.

3)Popular Music Genre: Rock

4)Top Rock Artist : AC/DC

5)Top invoice generator city: Prague emerges as the city with the highest total invoice sum. Other top contributors include Mountain View, London, and Berlin, making them potential candidates for a promotional festival, with Prague being the first choice as top two customers who spend the most money also belong to Prague.

Thank you!