SQL PROJECT

MUSIC STORE ANALYSIS

BY: ROOPAM TEWARI

OBJECTIVE

1.CUSTOMER BEHAVIOUR

2. SALES PERFORMANCE

3.GENRE TREND ANALYSIS

4.SPENDING BEHAVIOUR INSIGHTS

5.MARKET PENETRATION

We expect to gain actionable insights that will drive improved sales performance, optimized inventory management, and more targeted marketing strategies. This will result in better customer understanding, increased engagement, and enhanced operational efficiency. Ultimately, these outcomes will support data-driven decision-making and contribute to the store's growth and success.

DATABASE SCHEMA

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

EASY

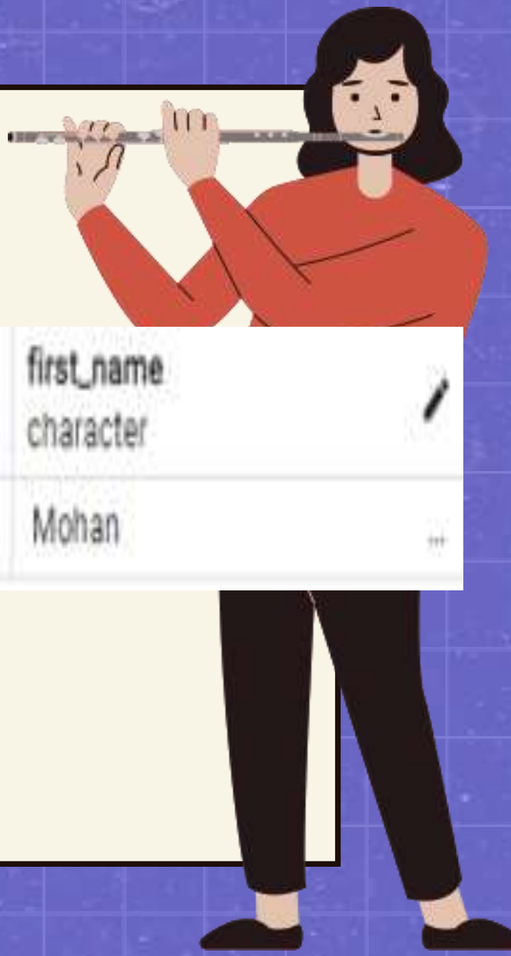
Q1. Who is the senior most employee based on job title?

QUERY

```
SELECT *
FROM employee
ORDER BY levels DESC
LIMIT 1;
```

OUTPUT

employee_id [PK] character varying (50)	last_name character	first_name character
9	Madan	Mohan



EASY

Q2.Which countries have the most Invoices?

QUERY

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

OUTPUT

	most_invoices 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

EASY

Q3. What are top 3 values of total invoice?

QUERY

```
SELECT total FROM invoice
ORDER BY total DESC
LIMIT 3
```

OUTPUT

	total double precision
1	23.759999999999998
2	19.8
3	19.8

EASY

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 invoice_total
FROM invoice
GROUP BY billing_city
ORDER BY invoice_total DESC
```

OUTPUT

	billing_city character varying (30)	invoice_total double precision
1	Prague	273.24000000000007
2	Mountain View	169.29
3	London	166.32
4	Berlin	158.4
5	Paris	151.47
6	São Paulo	129.69
7	Dublin	114.83000000000007

EASY

Q5. 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,C.first_name,c.last_name,SUM(i.total)AS TOTAL_AMOUNT
FROM customer AS c
INNER JOIN invoice AS i
ON c.customer_id=i.customer_id
GROUP BY c.customer_id,c.first_name,c.last_name
ORDER BY TOTAL_AMOUNT DESC
LIMIT 1
```

OUTPUT

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

MODERATE

Q1. 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 C.email,C.first_name,C.last_name
FROM customer AS C
JOIN invoice AS I ON C.customer_id=I.customer_id
JOIN invoice_line AS IL ON I.invoice_id=IL.invoice_id
JOIN track AS T ON IL.track_id=T.track_id
JOIN genre AS GR ON T.genre_id=GR.genre_id
WHERE GR.name ='Rock'
ORDER BY C.email ASC;
```

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

MODERATE

Q2. 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 ar.name AS ARTIST_NAME,COUNT(track_id)AS TOTAL_TRACK_COUNT
FROM artist AS ar
JOIN album AS a ON ar.artist_id=a.artist_id
JOIN track AS tr ON a.album_id=tr.album_id
JOIN genre AS g ON tr.genre_id=g.genre_id
WHERE g.name ='Rock'
GROUP BY ARTIST_NAME
ORDER BY TOTAL_TRACK_COUNT DESC
LIMIT 10;
```

OUTPUT

	artist_name character varying (120)	total_track_count bigint
1	Led Zeppelin	114
2	U2	112
3	Deep Purple	92
4	Iron Maiden	81
5	Pearl Jam	54
6	Van Halen	52

Total rows: 10 of 10 Query complete 00:00:00.150

MODERATE

Q3. . 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
FROM track
WHERE milliseconds >(SELECT AVG(milliseconds)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

ADVANCE

Q1. Find how much amount spent by each customer on artists? Write a query to return customer name, artist name and total spent

QUERY

```
SELECT c.first_name,c.last_name,ar.name AS artist_name,
SUM(il.unit_price*il.quantity)AS total_spent
FROM customer AS c
JOIN invoice AS i ON c.customer_id=i.customer_id
JOIN invoice_line AS il ON i.invoice_id=il.invoice_id
JOIN track AS tr ON il.track_id=tr.track_id
JOIN album AS a ON tr.album_id=a.album_id
JOIN artist AS ar ON a.artist_id=ar.artist_id
GROUP BY c.first_name,c.last_name,ar.name
ORDER BY total_spent DESC
```

OUTPUT

	first_name character	last_name character	artist_name character varying (120)	total_spent double precision
1	Hugh	O'Reilly	Queen	27.719999999999985
2	Wyatt	Girard	Frank Sinatra	23.759999999999999
3	Aaron	Mitchell	James Brown	19.799999999999997
4	François	Tremblay	The Who	19.799999999999997

ADVANCE

Q2. 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 genre_purchases AS(
    SELECT C.country,G.name AS genre,COUNT(IL.Quantity)AS purchases,
    ROW_NUMBER()OVER(PARTITION BY C.country ORDER BY COUNT(IL.Quantity)DESC)AS ROWNUMBER
FROM invoice_line AS IL
JOIN invoice AS I ON I.invoice_id=IL.invoice_id
JOIN Customer AS C ON C.customer_id=I.customer_id
JOIN track AS TR ON IL.track_id=TR.track_id
JOIN genre AS G ON TR.genre_id=G.genre_id
GROUP BY C.country,G.name
)
SELECT country,genre,purchases
FROM genre_purchases
WHERE ROWNUMBER=1
ORDER BY country,purchases DESC,genre
```

OUTPUT

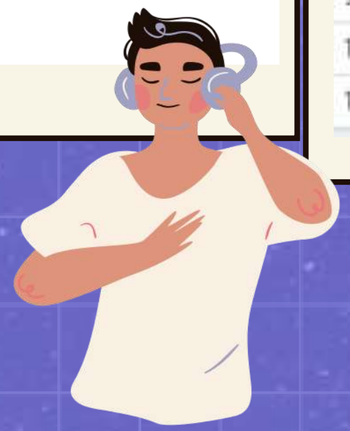
	country character varying (50)	genre character varying (120)	purchases bigint
1	Argentina	Alternative & Punk	17
2	Australia	Rock	34
3	Austria	Rock	40
4	Belgium	Rock	26
5	Brazil	Rock	205
6	Canada	Rock	333
7	Chile	Rock	61
8	Czech Republic	Rock	143
9	Denmark	Rock	24

ADVANCE

Q3. 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

```
WITH country_customer_spending AS(
    SELECT C.first_name,C.last_name AS CUSTOMER_NAME,billing_country AS COUNTRY,
    SUM(total)AS TOTAL_SPENT,
    ROW_NUMBER()OVER(PARTITION BY billing_country ORDER BY SUM(total)DESC)AS ROWNO
FROM invoice AS I
JOIN Customer AS C ON I.customer_id=C.customer_id
GROUP BY C.first_name,C.last_name,I.billing_country
ORDER BY COUNTRY ASC, TOTAL_SPENT DESC
)
SELECT * FROM country_customer_spending
WHERE ROWNO=1
```



OUTPUT

	first_name character	customer_name character	country character varying (30)	total_spent double precision	rowno bigint
1	Diego	Gutiérrez	Argentina	39.6	1
2	Mark	Taylor	Australia	81.18	1
3	Astrid	Gruber	Austria	69.3	1
4	Daan	Peeters	Belgium	60.38999999999999	1
5	Luís	Gonçalves	Brazil	108.89999999999998	1
6	François	Tremblay	Canada	99.99	1
7	Luis	Rojas	Chile	97.02000000000001	1
8	R	Madhav	Czech Republic	144.54000000000002	1
9	Kara	Nielsen	Denmark	37.61999999999999	1
10	Terhi	Hämäläinen	Finland	79.2	1
11	Wyatt	Girard	France	99.99	1

THANK YOU