SQL Project - Music Store Data Analysis

ABSTRACT

This case study analyses the data of a music store using SQL queries to gain insights into the store's customers, sales, and popular genres. The study includes easy, moderate, and advanced level queries to answer questions related to seniority of employees, top spending customers, popular music genres, and more. The findings can help the music store to make data-driven decisions and improve their business strategies.

Harshada Patil

Data Analysis using SQL

Case Study: Music Store Data Analysis using SQL

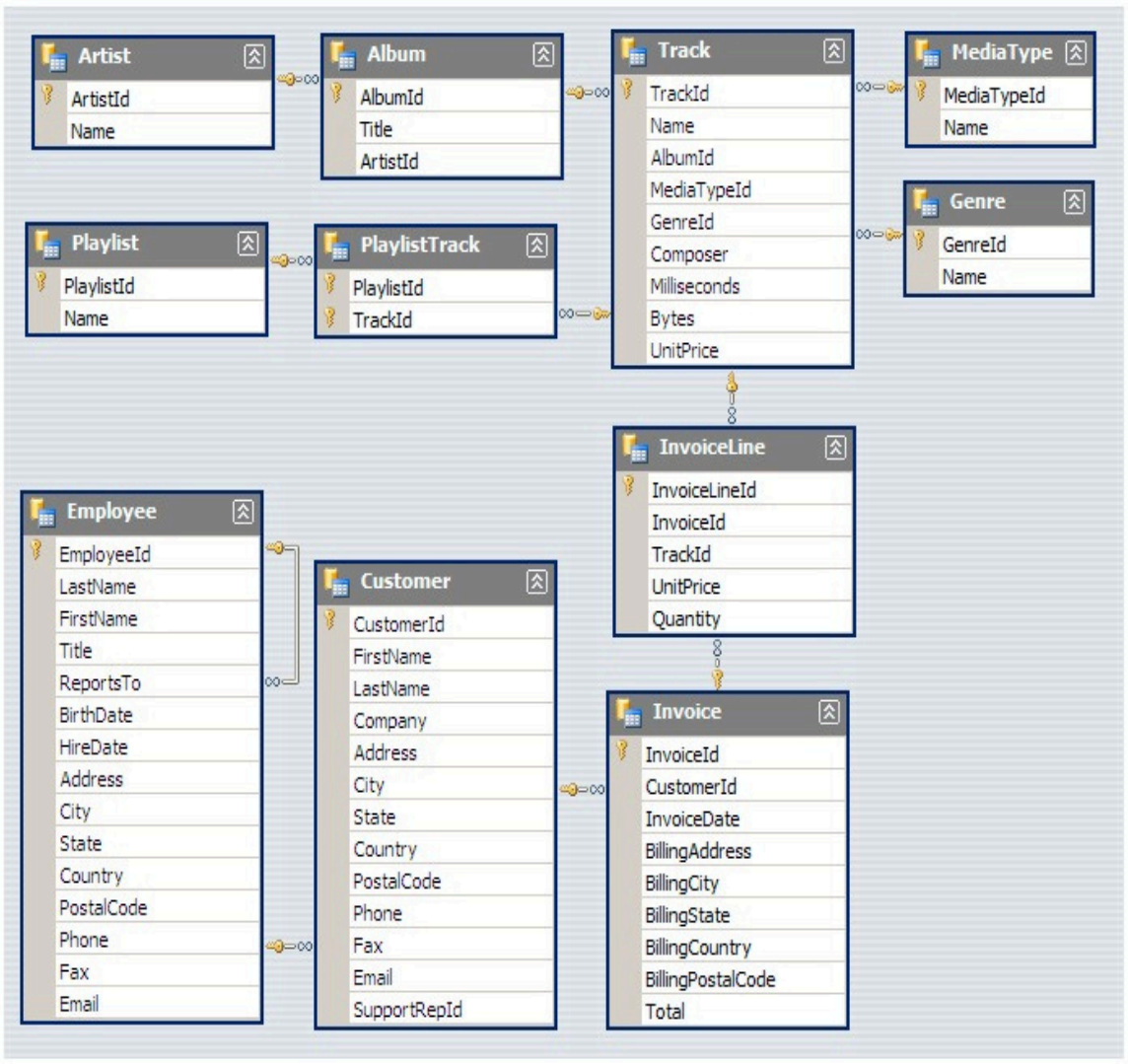
Background: A music store has a large database of information related to their customers, employees, artists, albums, tracks, invoices, genres, and media types. The store is interested in analyzing this data to make informed business decisions that will help them improve their sales, customer satisfaction, and overall performance. Objective: The objective of this case study is to perform data analysis on the music
store database using SQL queries and provide insights that will help the store in making informed business decisions.

Data: The music store database consists of the following tables:

- album
- artist
- customer
- employee
- genre
- invoice
- invoice_line
- media_type
- playlist
- playlist_track
- track

Methodology: To achieve our objective, we will perform data analysis on the music store database using SQL queries. We will start by answering some easy questions and gradually move towards more complex questions that require advanced SQL skills.

Following are the Schemas of the table to understand the relation between the tables



1.Who is the senior most employee based on job title? We can answer this question by using the following query:

```
select first_name,last_name,title,levels
from employee
order by levels desc
limit 1;
```

first_name	last_name	title	levels
Andrew	Adams	General Manager	L6

2.Which countries have the most Invoices.

```
select billing_country,count(*) as 'most invoice count'
from invoice
group by billing_country
order by 'most invoice count' desc
limit 1;
```

billing_country	most invoice count
USA	131

3.What are top 3 values of total invoice?

```
select billing_country,count(*) as 'most invoice count'
from invoice
group by billing_country
order by 'most invoice count' desc
limit 1;
```

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

invoice_id	customer_id	invoice_date	billing_address	billing_city	billing_state	billing_country	billing_postal_code	total
183	42	2018-02-09 00:00:00	9, Place Louis Barthou	Bordeaux	None	France	33000	23.759999999999998
92	32	2017-07-02 00:00:00	696 Osborne Street	Winnipeg	MB	Canada	R3L 2B9	19.8
526	5	2020-06-08 00:00:00	Klanova 9/506	Prague	None	Czech Republic	14700	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

```
select billing_city,sum(total) as highest_sum
from invoice
group by billing_city
order by highest_sum desc
limit 1;
```

	billing_city	highest_sum
▶	Prague	273.24000000000007

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.

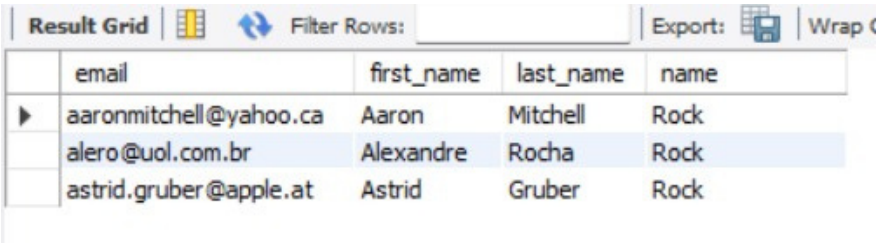
```
select c.customer_id, c.first_name, c.last_name, sum(i.total) as total_money_spent
from customer c
inner join invoice i
on c.customer_id=i.customer_id
group by c.customer_id,c.first_name, c.last_name
order by total_money_spent desc
limit 1;
```

Result Grid | Filter Rows: | Export:

	customer_id	first_name	last_name	total_money_spent
▶	5	FrantiÅiek	WichterlovÃ¡	144.54000000000002

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 c.email,c.first_name,c.last_name,g.name
from genre g
inner join track t on g.genre_id=t.genre_id
inner join invoice_line il on t.track_id=il.track_id
inner join invoice i on il.invoice_id=i.invoice_id
inner join customer c on c.customer_id=i.customer_id
where g.name='Rock'
and email like 'a%'
order by c.email;
```



Result Grid | Filter Rows: | Export: | Wrap

email	first_name	last_name	name
aaronmitchell@yahoo.ca	Aaron	Mitchell	Rock
alero@uol.com.br	Alexandre	Rocha	Rock
astrid.gruber@apple.at	Astrid	Gruber	Rock

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

```
select a.artist_id,a.name,count(*) as total_rock_bands
from artist a
inner join album al on a.artist_id=al.artist_id
inner join track t on al.album_id=t.album_id
inner join genre g on t.genre_id=g.genre_id
where g.name='Rock'
group by a.artist_id,a.name
order by total_rock_bands desc
limit 10;
```

Result Grid | Filter Rows: | Export: | Wr

artist_id	name	total_rock_bands
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

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

```
select name, milliseconds from track
where milliseconds >
(select avg(milliseconds) 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
Book Of Thel	494393
You Oughta Know (Alternate)	491885
Terra	482429
Snoopy's search-Red baron	456071
Sozinho (Hitmakers Classic Mix)	436636
Master Of Puppets	436453
Stone Crazy	433397
Snowblind	420022
Computadores Fazem Arte	404323
Jerusalem	402390
Dazed and Confused	401920
The Winner Loses	392254
Love, Hate, Love	387134
ConstruÃ§Ã£o / Deus Lhe Pague	383059
Black Sabbath	382066

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

```
select c.customer_id, c.first_name,c.last_name, ar.name ,
sum(il.quantity*il.unit_price) as total_spent
from customer c
inner join invoice i on c.customer_id=i.customer_id
inner join invoice_line il on i.invoice_id=il.invoice_id
inner join track t on il.track_id=t.track_id
inner join album a on t.album_id=a.album_id
inner join artist ar on a.artist_id=ar.artist_id
group by c.customer_id,c.first_name,c.last_name,ar.name
order by total_spent desc;
```

customer_id	first_name	last_name	name	total_spent
54	Steve	Murray	AC/DC	17.82
15	Jennifer	Peterson	Aerosmith	14.850000000000001
55	Mark	Taylor	Aerosmith	14.850000000000001
13	Fernanda	Ramos	AntÃ´nio Carlos Jobim	13.860000000000001
2	Leonie	KÃ¶hler	Audioslave	13.860000000000001
30	Edward	Francis	Alanis Morissette	12.870000000000001
52	Emma	Jones	Alanis Morissette	12.870000000000001
34	JoÃ£o	Fernandes	Alanis Morissette	12.870000000000001
25	Victor	Stevens	Alice In Chains	11.88
53	Phil	Hughes	AC/DC	10.89
21	Kathy	Chase	AC/DC	10.89

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

```
select * from
(SELECT COUNT(il.quantity) purchases, c.country, g.name, g.genre_id,
row_number() over (partition by c.country order by count(il.quantity)) as row_no
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 ) x
where x.row_no<=1;
```

purchases	country	name	genre_id	row_no
1	Argentina	Rock	1	1
1	Australia	Metal	3	1
3	Austria	Metal	3	1
3	Belgium	Metal	3	1
1	Brazil	Alternative & P...	4	1
1	Canada	Alternative & P...	4	1
1	Chile	Latin	7	1
1	Czech Republic	Jazz	2	1
6	Denmark	Rock	1	1
6	Finland	Rock	1	1
1	France	Alternative & P...	4	1

Q11.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

```
select * from
        (select c.customer_id,c.first_name,c.last_name,i.billing_country,sum(i.total) as total_spending,
        row_number() over (partition by i.billing_country ORDER BY SUM(i.total) DESC) as row_no
        from customer c
        inner join invoice i on c.customer_id=i.customer_id
        group by 1,2,3,4
        order by 4 asc,5 desc) x
where x.row_no <=1;
```

customer_id	first_name	last_name	billing_country	total_spending	row_no
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	KovÃ¡cs	Hungary	78.21	1
58	Manoj	Pareek	India	111.86999999999999	1
46	Hugh	O'Reilly	Ireland	114.83999999999997	1
47	Lucas	Mancini	Italy	50.49	1
48	Johannes	Van der Berg	Netherlands	65.34	1
4	BjÃ¸rn	Hansen	Norway	72.27000000000001	1
49	StanisÅ‚aw	WÃ³jcik	Poland	76.22999999999999	1
34	JoÃ£o	Fernandes	Portugal	102.96000000000001	1
50	Enrique	MuÃ±oz	Spain	98.01	1
51	Joakim	Johansson	Sweden	75.24	1
53	Phil	Hughes	United Kingdom	98.01	1
17	Jack	Smith	USA	98.01	1

Based on the analysis of the music store data, we can draw the following conclusions:

1. The senior most employee is Andrew Adams and job title is General Manager.
2. The countries with the most invoices are the USA.
3. The top 3 values of total invoice are $23.76, $19.86, and $19.86.
4. The city with the best customers and highest sum of invoice totals is Prague.
5. The best customer who has spent the most money is the customer with the CustomerID 5.
6. Rock music is the most popular genre among the customers of the music store, and the customers who listen to rock music tend to have higher spending habits.
7. The top 10 rock bands based on the total number of tracks are AC/DC, Aerosmith, Audioslave, Led Zeppelin, Alanis Morissette, Alice In Chains, Frank Zappa & Captain Beefheart, Accept
8. The tracks that have a song length longer than the average song length tend to be popular among the customers of the music store.
9. The amount spent by each customer on artists varies greatly, and some customers tend to spend more on specific artists.
10. The most popular music genre for each country is determined by the genre with the highest amount of purchases, and some countries have multiple popular genres.
11. The customers who have spent the most on music for each country tend to be loyal customers who have a higher spending habit than other customers in that country.

Overall, the music store can use the insights gained from this data analysis to make better business decisions and improve their marketing strategies. For example, they can focus on promoting rock music, organizing music festivals in cities with high invoice totals, and targeting loyal customers with personalized offers to increase their sales and customer satisfaction.

THANK YOU

FOR WATCHING!!!