Music Store Analysis

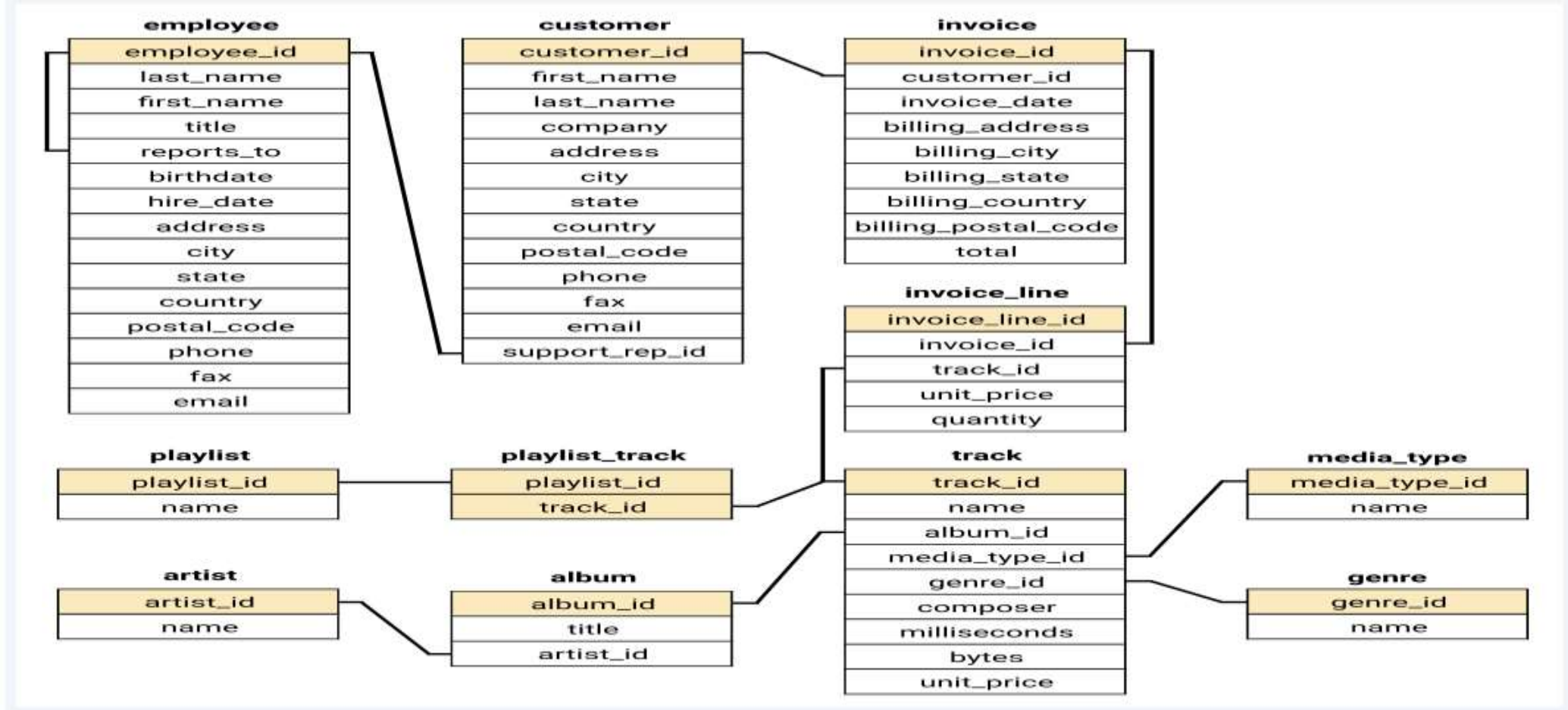
SQL Project

-Drishti Sharma

Objective

- The project aims to analyze a digital music store database using SQL, providing stakeholders with valuable insights for decision-making.
- Through SQL queries, it addresses questions regarding geographical growth, purchase power, revenue, genre performance, and music band popularity.
- The analysis offers actionable recommendations to drive business growth and optimize performance.

Database Schema



Level Of Queries

EASY

- Select , Group by
- Order By, Limit
- Desc/Asc

MODERATE

- Joins, Order by
- Group by, Limits

HARD

- CTE(Common table Expression)

Easy

Question 1. 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	title character varying (50)	reports_to character varying (30)	levels character varying (10)
9	Madan	Mohan ...	Senior General Manager	[null]	L7

Easy

Question 2. Which countries have the most Invoices?

- Query

```
select billing_country, count(*) as invoice_count
from invoice
group by billing_country
order by count(*) desc
```

- Output

billing_country character varying (30)	invoice_count bigint
USA	131
Canada	76
Brazil	61
France	50
Germany	41
Czech Republic	30
Portugal	29
United Kingdom	28
India	21
Chile	13

Easy

Question 3. What are top values of total invoice and from which country?

- Query

```
select total from invoice
order by total desc
limit 3
```

- Output

total double precision
23.759999999999998
19.8
19.8

Easy

Question 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

- Query

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

- Output

billing_city character varying (30)	invoice_total double precision
Prague	273.24000000000007

Easy

Question 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 customer.customer_id, customer.first_name, customer.last_name, sum(invoice.total) as total
from customer
Join invoice on customer.customer_id =invoice.customer_id
group by customer.customer_id
order by total desc
limit 1;
```

- Output

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

Moderate

Question 1. 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
join invoice on customer.customer_id = invoice.customer_id
join invoice_line on invoice.invoice_id = invoice_line.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
```

- Output

email character varying (50)	first_name character	last_name character
aaronmitchell@yahoo.ca	Aaron	Mitchell
alero@uol.com.br	Alexandre	Rocha
astrid.gruber@apple.at	Astrid	Gruber
bjorn.hansen@yahoo.no	Bjørn	Hansen
camille.bernard@yahoo.fr	Camille	Bernard
daan_peeters@apple.be	Daan	Peeters
diego.gutierrez@yahoo.ar	Diego	Gutiérrez
dmiller@comcast.com	Dan	Miller
dominiquelefebvre@gmail.c...	Dominique	Lefebvre

Moderate

Question 2. 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 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;
```

- Output

name character varying (120)	number_of_songs bigint
Led Zeppelin	114
U2	112
Deep Purple	92
Iron Maiden	81
Pearl Jam	54
Van Halen	52
Queen	45
The Rolling Stones	41
Creedence Clearwater Revival	40
Kiss	35

Moderate

Question 3. 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) as avg_track_length
    from track
)
order by milliseconds desc;
```

- Output

name character varying (150)	milliseconds integer
Occupation / Precipice	5286953
Through a Looking Glass	5088838
Greetings from Earth, Pt. 1	2960293
The Man With Nine Lives	2956998
Battlestar Galactica, Pt. 2	2956081
Battlestar Galactica, Pt. 1	2952702
Murder On the Rising Star	2935894
Battlestar Galactica, Pt. 3	2927802

Advance

Question 1. Find how much amount spent by each customer on artists? Write a query to return customer name, artist name and total spent

- Query

```
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;
```

- Output

customer_id integer	first_name character	last_name character	artist_name character varying (120)	amount_spent double precision
46	Hugh	O'Reilly	Queen	27.719999999999985
38	Niklas	Schröder	Queen	18.81
3	François	Tremblay	Queen	17.82
34	João	Fernandes	Queen	16.830000000000002
53	Phil	Hughes	Queen	11.88

Advance

Question :2. 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 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
```

- Output

purchases bigint	country character varying (50)	name character varying (120)	genre_id character varying (50)	rowno bigint
17	Argentina	Alternative & Punk	4	1
34	Australia	Rock	1	1
40	Austria	Rock	1	1
26	Belgium	Rock	1	1
205	Brazil	Rock	1	1
333	Canada	Rock	1	1

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 customer_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 customer_with_country where RowNo <= 1
```

- Output

customer_id integer	first_name character	last_name character	billing_country character varying (30)	total_spending double precision	rowno bigint
56	Diego	Gutiérrez	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	Luís	Gonçalves	Brazil	108.89999999999998	1
3	François	Tremblay	Canada	99.99	1

Thank you ☺