SQL PROJECT
MUSIC STORE ANALYSIS

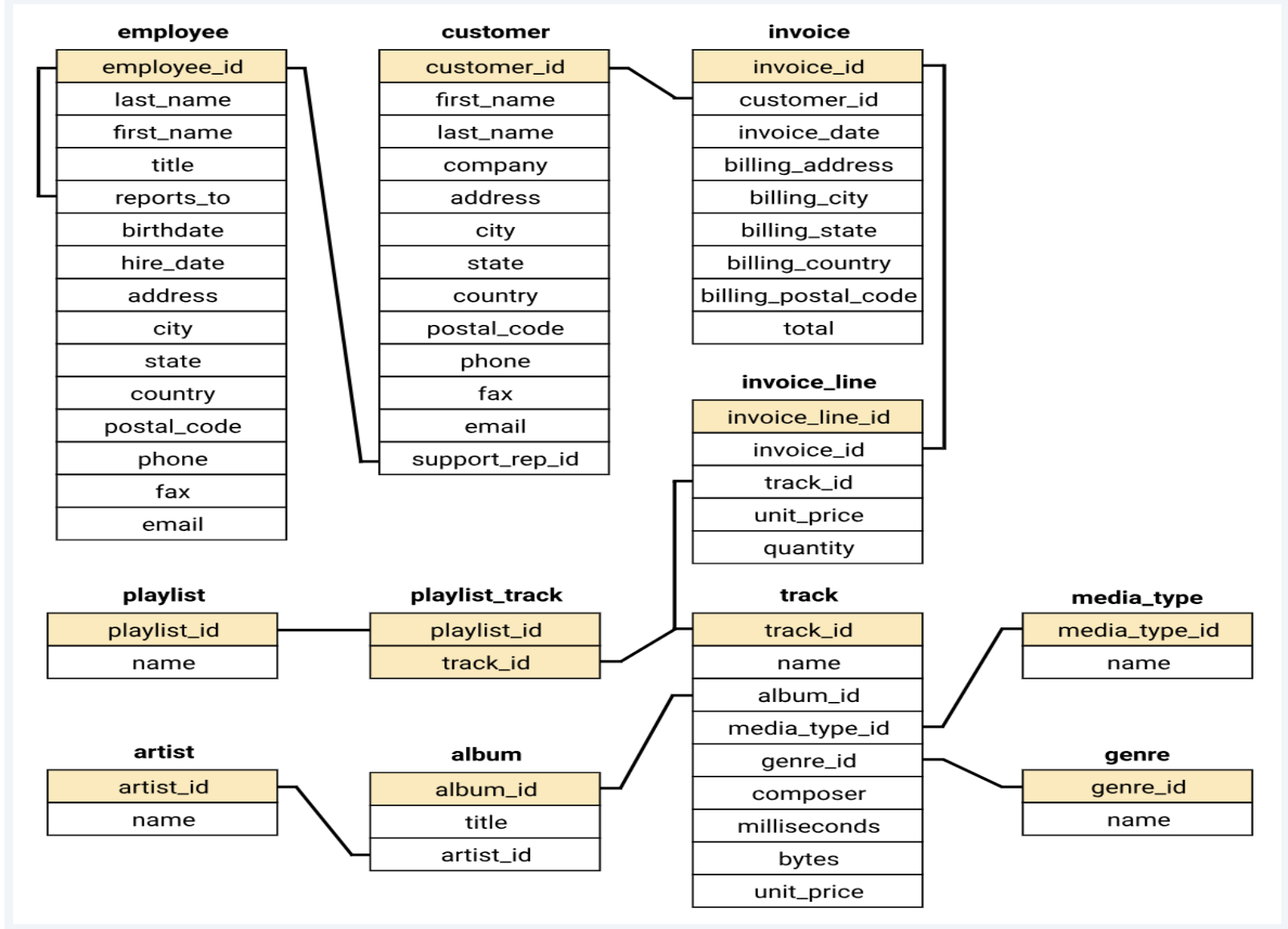
Project By: Prinshi Jha

TABLE OF CONTENT

Project Overview

- This SQL project focuses on analyzing a music database using multiple queries. It utilizes various tables to extract and evaluate important data, offering valuable insights that improve strategic decision-making and optimize marketing and sales strategies.
- Tool used:
- 1. PostgreSQL: Used as the primary database in a music store analysis project to store and manage data on customers, songs, albums, and sales, allowing for complex queries to gain insights into customer behavior and sales trends.
- 2. pgAdmin4: A management tool for interacting with the PostgreSQL database, enabling visualization of music store data, execution of SQL queries for analysis, and easy monitoring of data relationships and sales performance.

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

Basic Queries

1. Identify the most senior employee based on their job title.

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

	first_name character (50)	last_name character (50)	title character varying (50)
1	Mohan	Madan	Senior General Manager

2. Determine the country with the highest number of invoices and provide a count for each.

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

	invoice_count bigint	billing_country character varying (30)
1	131	USA
2	76	Canada
3	61	Brazil
4	50	France
5	41	Germany

3. Retrieve the top three values of total invoices.

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

	total double precision
1	23.759999999999998
2	19.8
3	19.8

4. Identify the city with the highest sum of invoice totals, as we plan a promotional Music Festival.

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

	billing_city character varying (30)	invoice_total double precision
1	Prague	273.24000000000007

5. Find the customer who has spent the most money.

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

	customer_id [PK] integer	first_name character (50)	last_name character (50)	total double precision
1	5	R ...	Madhav	144.54000000000002

6. Determine the average total for all invoices.

```
SELECT AVG(total) AS average_invoice_total
FROM invoice;
```

	average_invoice_total double precision
1	7.670081433224746

7. Which genre has the highest number of tracks in the database.

```
SELECT genre.name, COUNT(track.track_id) AS track_count
FROM genre
JOIN track ON genre.genre_id = track.genre_id
GROUP BY genre.name
ORDER BY track_count DESC
LIMIT 1;
```

	name character varying (120)	track_count bigint
1	Rock	1297

Intermediate Queries

1. Retrieve the email, first name and last name of all Rock Music listener, ordered alphabetically by email.

```sql
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 = 'Rock'
       )
ORDER BY email;
```

	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

2. Identify the top 10 rock bands based on the total number of tracks they have written.

```
SELECT artist.artist_id, artist.name,
COUNT(artist.artist_id) as number_of_songs
FROM track
JOIN album ON track.album_id = album.album_id
JOIN artist ON album.artist_id = artist.artist_id
JOIN genre ON track.genre_id = genre.genre_id
WHERE genre.name = 'Rock'
GROUP BY artist.artist_id
ORDER BY Number_of_songs DESC
LIMIT 10;
```

	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

3. Retrieve track names and milliseconds for tracks longer than the average song length.

```
SELECT name, milliseconds
FROM track
Where milliseconds > (SELECT avg(milliseconds) FROM track)
Order by Milliseconds DESC;
```

	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

4. List all artists with albums containing more than 20 tracks.

```sql
SELECT artist.name, COUNT(track.track_id) AS track_count
FROM artist
JOIN album ON artist.artist_id = album.artist_id
JOIN track ON album.album_id = track.album_id
GROUP BY artist.name
HAVING COUNT(track.track_id) > 20
ORDER BY track_count DESC;
```

	name character varying (120)	track_count bigint
1	Iron Maiden	213
2	U2	135
3	Led Zeppelin	114
4	Metallica	112
5	Deep Purple	92
6	Lost	92
7	Pearl Jam	67
8	Lenny Kravitz	57

Advanced Queries

1. List the top 5 customers who have made the highest total purchases.

```sql
SELECT customer.customer_id, customer.first_name,
customer.last_name, SUM(invoice.total) AS total_purchases
FROM customer
JOIN invoice ON customer.customer_id = invoice.customer_id
GROUP BY customer.customer_id, customer.first_name, customer.last_name
ORDER BY total_purchases DESC
LIMIT 5;
```

	customer_id [PK] integer	first_name character (50)	last_name character (50)	total_purchases double precision
1	5	R	Madhav	144.54000000000002
2	6	Helena	Holý	128.7
3	46	Hugh	O'Reilly	114.83999999999997
4	58	Manoj	Pareek	111.86999999999999
5	1	Luís	Gonçalves	108.89999999999998

2. Identify the top 3 countries that have the highest average invoice total. Return the country name and the average invoice total.

```sql
SELECT billing_country, AVG(total) AS average_invoice_total
FROM invoice
GROUP BY billing_country
ORDER BY average_invoice_total DESC
LIMIT 3;
```

	billing_country character varying (30)	average_invoice_total double precision
1	Czech Republic	9.108000000000002
2	Spain	8.91
3	Ireland	8.833846153846151

3. Find the artists who have tracks in multiple genres. Return the artist name and the count of distinct genres.

```sql
SELECT artist.name AS artist_name,
COUNT(DISTINCT genre.genre_id) AS distinct_genre_count
FROM artist
JOIN album ON artist.artist_id = album.artist_id
JOIN track ON album.album_id = track.album_id
JOIN genre ON track.genre_id = genre.genre_id
GROUP BY artist_name
HAVING COUNT(DISTINCT genre.genre_id) > 1
ORDER BY distinct_genre_count DESC;
```

	artist_name character varying (120)	distinct_genre_count bigint
1	Iron Maiden	4
2	Audioslave	3
3	Battlestar Galactica	3
4	Various Artists	3
5	Gilberto Gil	3

Key Findings

- Identified key markets and high-value customers through analysis of invoices, spending, and top transactions.
- Pinpointed the best cities and countries for promotional events based on invoice totals and spending patterns.
- Highlighted the top 10 rock artists and provided insights into rock music listeners for targeted promotions.
- Analyzed customer preferences by tracking spending on artists and popular music genres across countries.
- Curated playlists and marketing strategies based on insights into song length and regional genre preferences.

Thank You