SQL PROJECT

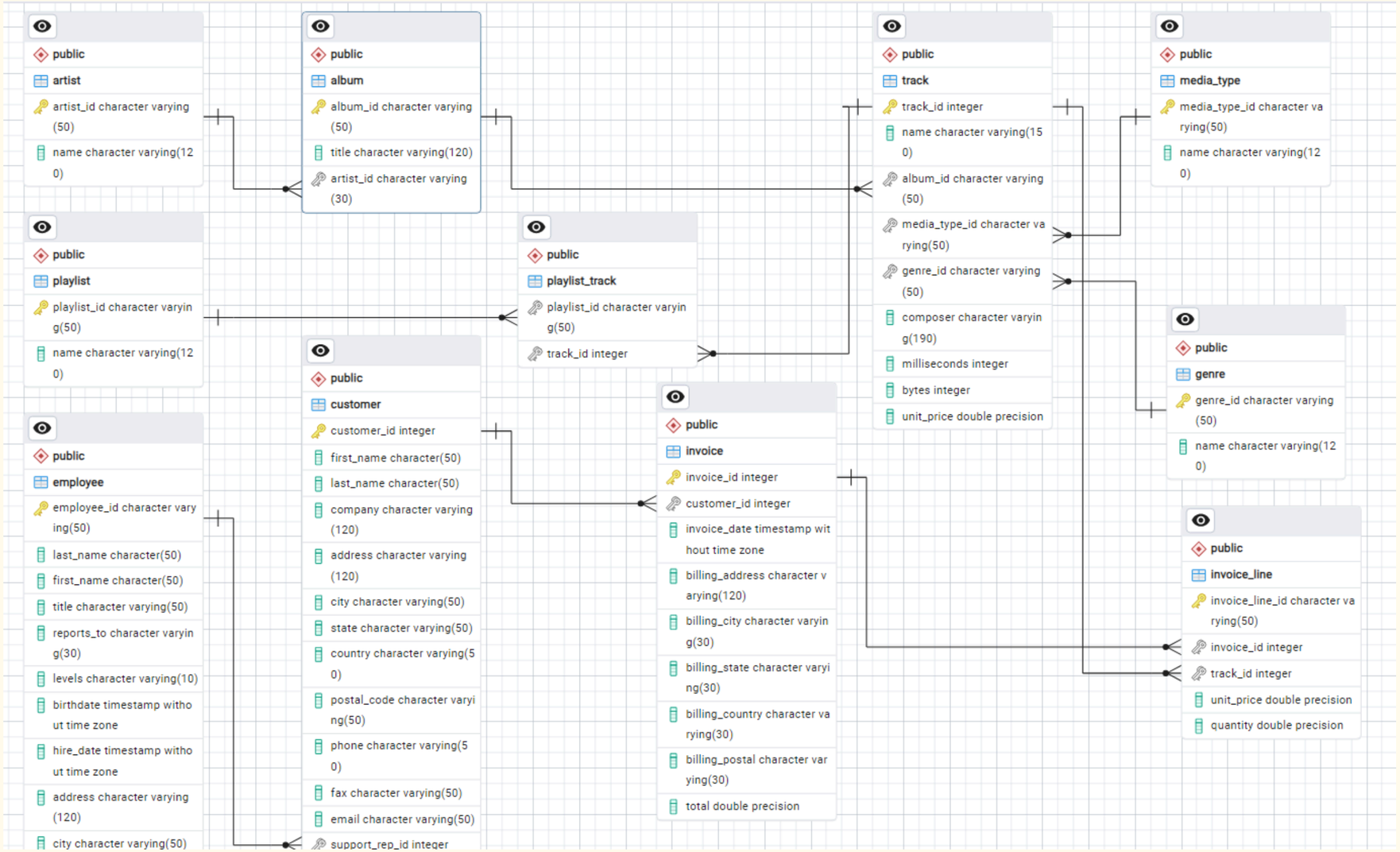
MUSIC STORE ANALYSIS

Presented by: ABHISHEK SINGH RAO

Objective :

Achieve substainable Business growth for the music store by identifying key trends, customer behaviour and product performance through data analysis.

Database Schema



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

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

	first_name character	last_name character	title character varying (50)	levels character varying (10)
1	Mohan	Madan	Senior General Manager	L7

Q2. Which countries have the most Invoices?

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

	billing_country character varying (30)	invoice_count bigint
1	USA	131
2	Canada	76
3	Brazil	61
4	France	50
5	Germany	41
6	Czech Republic	30
7	Portugal	29
8	United Kingdom	28
9	India	21
10	Chile	13

Total rows: 24 of 24 | Query complete 00:00:00.113

Q3. What are top 3 values of total invoice?

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

	total double precision
1	23.759999999999998
2	19.8
3	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

```
SELECT
    billing_city,
    SUM(total) AS invoice_total
FROM invoice
GROUP BY billing_city
ORDER BY SUM(total) DESC
```

	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	114.83999999999997
8	Delhi	111.86999999999999
9	São José dos Campos	108.89999999999998
10	Brasília	106.91999999999999
11	Lisbon	102.96000000000001
12	Bordeaux	99.99
13	Montréal	99.99
14	Madrid	98.01

Total rows: 53 of 53 Query complete 00:00:00.10

Q5. Who is the best customer? The customer who has spent the most money will be declared the best customer.

```sql
SELECT
    c.customer_id,
    c.first_name,
    c.last_name,
    SUM(i.total) AS total_spending
FROM customer AS c
LEFT JOIN invoice AS i
ON c.customer_id = i.customer_id
GROUP BY c.customer_id, c.first_name, c.last_name
ORDER BY SUM(i.total) DESC
LIMIT 1
```

	customer_id [PK] integer	first_name character	last_name character	total_spending double precision
1	5	R ...	Madhav	144.54000000000002

Q6. 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
FROM genre AS g
JOIN track AS t
    ON g.genre_id = t.genre_id
JOIN invoice_line AS il
    ON t.track_id = il.track_id
JOIN invoice AS i
    ON il.invoice_id = i.invoice_id
JOIN customer AS c
    ON i.customer_id = c.customer_id
WHERE g.name = 'Rock'
ORDER BY c.email
```

	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
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
11	eduardo@woodstock.com.br	Eduardo	Martins
12	ellie.sullivan@shaw.ca	Ellie	Sullivan
13	emma_jones@hotmail.com	Emma	Jones

Total rows: 59 of 59 Query complete 00:00:00.130

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.

```sql
SELECT
    art.artist_id,
    art.name,
    COUNT(tck.*)
FROM artist AS art
JOIN album AS alb
    ON art.artist_id = alb.artist_id
JOIN track AS tck
    ON alb.album_id = tck.album_id
JOIN genre AS gnr
    ON tck.genre_id = gnr.genre_id
WHERE gnr.name = 'Rock'
GROUP BY art.artist_id, art.name
ORDER BY COUNT(tck.*) DESC
LIMIT 10
```

	artist_id [PK] character varying (50)	name character varying (120)	count 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

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.

```sql
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
7	Murder On the Rising Star	2935894
8	Battlestar Galactica, Pt. 3	2927802
9	Take the Celestra	2927677
10	Fire In Space	2926593
11	The Long Patrol	2925008
12	The Magnificent Warriors	2924716
13	The Living Legend, Pt. 1	2924507
14	The Gun On Ice Planet Zero, Pt. 2	2924341
15	The Hand of God	2924007
16	Experiment In Terra	2923548

Total rows: 494 of 494 Query complete 00:00:00.

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

```sql
WITH best_selling_artist AS (
    SELECT art.artist_id, art.name, SUM(il.unit_price * il.quantity) AS total_amount
    FROM artist AS art
    JOIN album AS abm ON art.artist_id = abm.artist_id
    JOIN track AS trk ON abm.album_id = trk.album_id
    JOIN invoice_line AS il ON trk.track_id = il.track_id
    GROUP BY art.artist_id, art.name
    ORDER BY SUM(il.unit_price * il.quantity) DESC
    LIMIT 1
)

SELECT c.customer_id, c.first_name, c.last_name, bsa.name,
    SUM(inl.unit_price * inl.quantity)
FROM customer AS c
JOIN invoice AS i ON c.customer_id = i.customer_id
JOIN invoice_line AS inl ON i.invoice_id = inl.invoice_id
JOIN track AS tck ON inl.track_id = tck.track_id
JOIN album AS alb ON tck.album_id = alb.album_id
JOIN best_selling_artist AS bsa ON alb.artist_id = bsa.artist_id
GROUP BY c.customer_id, c.first_name, c.last_name, bsa.name
ORDER BY SUM(inl.unit_price * inl.quantity) DESC;
```

	customer_id integer	first_name character	last_name character	name character varyin	sum double precision
1	46	Hugh	O'Reilly	Queen	27.719999999999985
2	38	Niklas	Schröder	Queen	18.81
3	3	François	Tremblay	Queen	17.82
4	34	João	Fernandes	Queen	16.830000000000002
5	53	Phil	Hughes	Queen	11.88
6	41	Marc	Dubois	Queen	11.88
7	47	Lucas	Mancini	Queen	10.89
8	33	Ellie	Sullivan	Queen	10.89
9	20	Dan	Miller	Queen	3.96
10	5	R	Madhav	Queen	3.96
11	23	John	Gordon	Queen	2.9699999999999998
12	54	Steve	Murray	Queen	2.9699999999999998
13	31	Martha	Silk	Queen	2.9699999999999998
14	16	Frank	Harris	Queen	1.98

Total rows: 43 of 43 | Query complete 00:00:00.048

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.

```sql
WITH popular_music_genre AS
(
    SELECT i.billing_country,
        g.genre_id,
        g.name,
        COUNT(inl.*) AS purchase,
        ROW_NUMBER() OVER(PARTITION BY i.billing_country ORDER BY COUNT(inl.*) DESC) AS rnk
    FROM genre as g
    JOIN track as t
    ON g.genre_id = t.genre_id
    JOIN invoice_line AS inl
    ON t.track_id = inl.track_id
    JOIN invoice AS i
    ON inl.invoice_id = i.invoice_id
    GROUP BY 1,2

)
SELECT
    billing_country,
    genre_id,
    name,
    purchase
FROM popular_music_genre
WHERE rnk = 1
```

	billing_country character varying (30)	genre_id character varying (50)	name character varying (120)	purchase bigint
1	Argentina	4	Alternative & Punk	17
2	Australia	1	Rock	34
3	Austria	1	Rock	40
4	Belgium	1	Rock	26
5	Brazil	1	Rock	205
6	Canada	1	Rock	333
7	Chile	1	Rock	61
8	Czech Republic	1	Rock	143
9	Denmark	1	Rock	24
10	Finland	1	Rock	46
11	France	1	Rock	211
12	Germany	1	Rock	194
13	Hungary	1	Rock	44

Total rows: 24 of 24 Query complete 00:00:00.157

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.

```
    WITH top_customer AS
    (
    SELECT
        i.billing_country,
        c.customer_id,
        c.first_name,
        c.last_name,
        SUM(total) AS total_spent,
        DENSE_RANK() OVER(PARTITION BY i.billing_country ORDER BY SUM(total) desc) as rnk
FROM customer AS c
JOIN invoice AS i
    ON c.customer_id = i.customer_id
GROUP BY 1,2,3
)
SELECT customer_id,
    first_name,
    last_name,
    billing_country,
    total_spent
FROM top_customer
WHERE rnk = 1
```

	customer_id integer	first_name character	last_name character	billing_country character varying (30)	total_spent double precision
1	56	Diego	Gutiérrez	Argentina	39.6
2	55	Mark	Taylor	Australia	81.18
3	7	Astrid	Gruber	Austria	69.3
4	8	Daan	Peeters	Belgium	60.38999999999999
5	1	Luís	Gonçalves	Brazil	108.89999999999998
6	3	François	Tremblay	Canada	99.99
7	57	Luis	Rojas	Chile	97.02000000000001
8	5	R	Madhav	Czech Republic	144.54000000000002
9	9	Kara	Nielsen	Denmark	37.61999999999999
10	44	Terhi	Hämäläinen	Finland	79.2
11	42	Wyatt	Girard	France	99.99
12	37	Fynn	Zimmermann	Germany	94.05000000000001
13	45	Ladislav	Kovács	Hungary	78.21
14	58	Manoj	Pareek	India	111.86999999999999
15	46	Hugh	O'Reilly	Ireland	114.83999999999997

Total rows: 24 of 24 | Query complete 00:00:00.130

Thank You

Abhishek Singh Rao