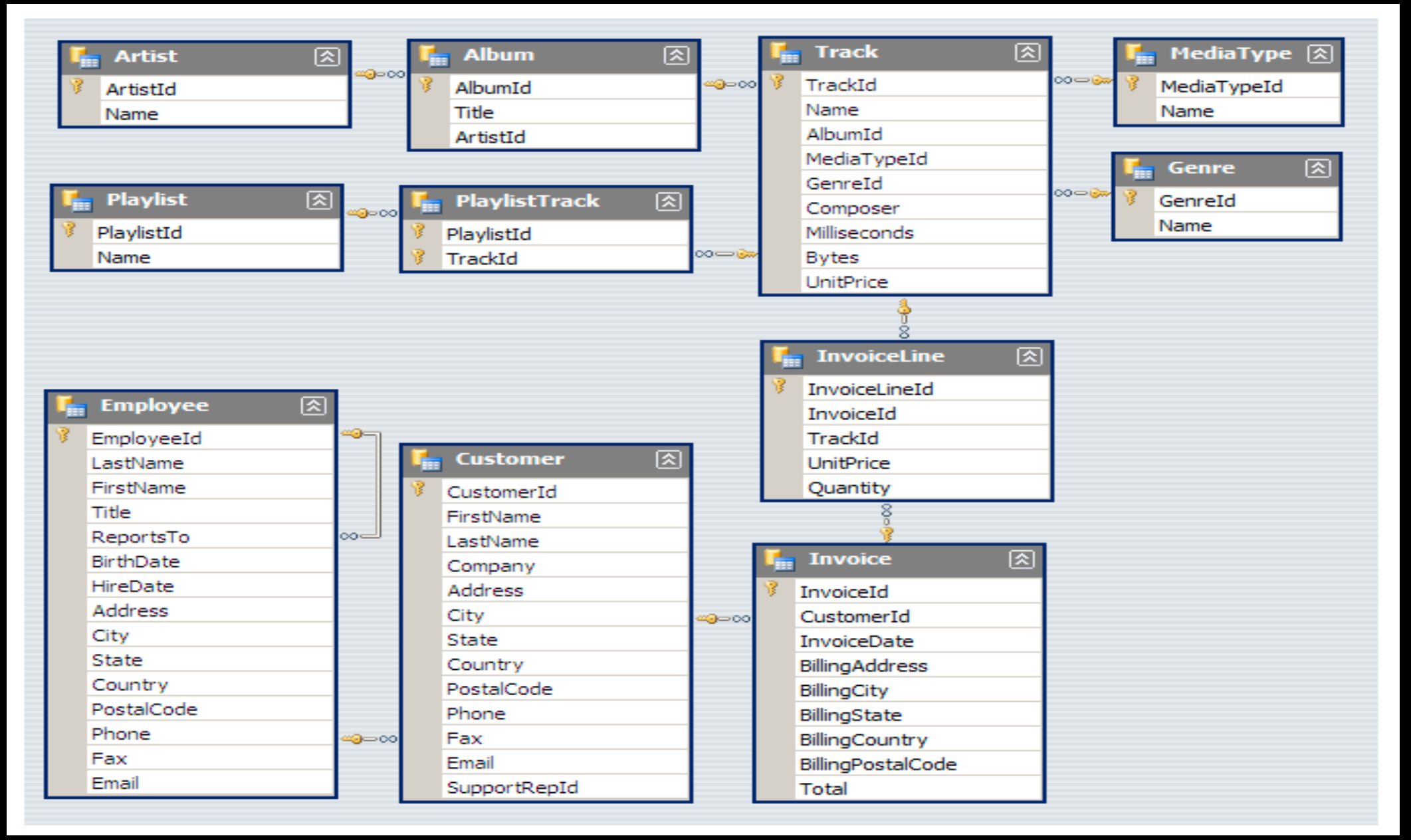

Music Store
Data Analysis Project
Using SQL
Music
Music
Music
Music
Music
Music
Music
Music
Music
Music
Music

Objective

- Analyse data related to sales, genres, and artists.
- Analyse sales data to understand the distribution of sales across different countries.
- Determine the popularity of music genres and Artist in each country.
- Analyse which genres are most favoured by customers in different geographic locations, which can inform inventory selection and marketing efforts.

Entity Relationship



SQL Analysis with their Output

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

Query | Query History

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

Select * from employee
ORDER BY levels desc
Limit 1
```

Scratch Pad

Messages | Notifications

Data Output

employee_id [PK] character varying (50)	last_name character	first_name character	title character varying (50)	reports_to character varying (30)	levels charact
9	Madan ...	Mohan ...	Senior General Manager	[null]	L7

SQL Analysis with their Output

Q2- Which countries have the most Invoices?

```
7   /*Which countries have the most Invoices?*/
8
9   select count(total)AS c , billing_country from invoice
10  Group by billing_country
11  Order by c desc
12
```

Messages Notifications

ata Output

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

SQL Analysis with their Output

Q3-What are top 3 values of total invoice?

```
13  /*3. What are top 3 values of total invoice?*/
14  select total from invoice
15  order by total desc
16  Limit 3
```

Messages Notifications

Data Output

total double precision
23.759999999999998
19.8
19.8

SQL Analysis with their Output

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

SQL Analysis with their Output

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

```
28  -- Who is the best customer? The customer who has spent the most money
29  -- will be declared the best customer.
30  -- Write a query that returns the person who has spent the most money.S
31  -- Select  * from customer
32
33  Select customer.customer_id, customer.first_name,customer.last_name,sum(invoice.total) as total
34  from customer
35  JOIN invoice ON customer.customer_id=invoice.customer_id
36  group by customer.customer_id
37  Order by total desc
38  limit 1
```

Messages Notifications

```
Successfully run. Total query runtime: 535 msec.
```

ata Output

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

SQL Analysis with their Output

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

```
43  Select Distinct first_name, last_name , Email
44  from customer
45  JOIN invoice ON invoice.invoice_id = customer.customer_id
46  JOIN invoice_line ON invoice_line.invoice_iD =invoice.invoice_id
47  WHERE track_id IN(
48      Select track_id from track
49      JOIN genre on track.genre_id = genre.genre_id
50      WHERE genre.name like 'Rock'
51  )
52  ORDER BY email
53  )
54
```

ata Output

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

SQL Analysis with their Output

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

```
58  Select artist.artist_id, artist.name,COUNT(artist.artist_id) as t_track
59  from track
60  JOIN album ON album.album_id = track.album_id
61  JOIN artist ON artist.artist_id = album.artist_id
62  JOIN genre ON genre.genre_id = track.genre_id
63  Where genre.name LIKE 'Rock'
64  Group by artist.artist_id
65  Order by t_track desc
66  Limit 10;
```

Messages Notifications

```
Successfully run. Total query runtime: 184 msec.
```

Data Output

artist_id [PK] character varying (50)	name character varying (120)	t_track bigint
22	Led Zeppelin	114
150	U2	112
58	Deep Purple	92
90	Iron Maiden	81
118	Pearl Jam	54
152	Van Halen	52

SQL Analysis with their Output

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

```
69  -- Return all the track names that have a song length longer than the average song length.
70  -- Return the Name and Milliseconds for each track. Order by the song length with the longest songs listed first
71
72  Select name , milliseconds FROM track
73  WHERE milliseconds > (
74          SELECT AVG(milliseconds) AS track_len
75          FROM track
76  )
77  Order by milliseconds desc;
78
79
```

Messages Notifications

```
Successfully run. Total query runtime: 200 msec.
494 rows affected.
```

Data 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

SQL Analysis with their Output

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 artist.artist_id AS artist_id , artist.name AS artist_fn,
    SUM(invoice_line.unit_price * invoice_line.Quantity) AS total_spent
    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
)
```

```sql
select c.customer_id , c.first_name,c.last_name,bsa.artist_fn,
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 al ON al.album_id=t.album_id
JOIN best_selling_artist as bsa ON bsa.artist_id=al.artist_id
Group by 1,2,3,4
Order by 5 desc;
```

Messages Notifications

```
Successfully run. Total query runtime: 148 msec.
43 rows affected.
```

Data Output

customer_id integer	first_name character	last_name character	artist_fn 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

SQL Analysis with their Output

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_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
```

Messages Notifications

Successfully run. Total query runtime: 137 msec.

ata 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

SQL Analysis with their Output

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

```sql
-- 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 Customter_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 Customter_with_country WHERE RowNo <= 1
```

ssages Notifications

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

Thank You