MUSIC STORE DATA ANALYSIS

Introduction :

The purpose of this project report is to present the results of the data analysis conducted on a music store's dataset. The data analysis aimed to gain insights into the performance and trends of the music store, analyse customer behaviour, identify popular music genres, and make data-driven recommendations to optimize the store's operations and enhance its revenue and customer satisfaction

I am using for making this project **pgadmin4, PoatgreSQL**

Project Objectives :

The main objectives of this data analysis project were:

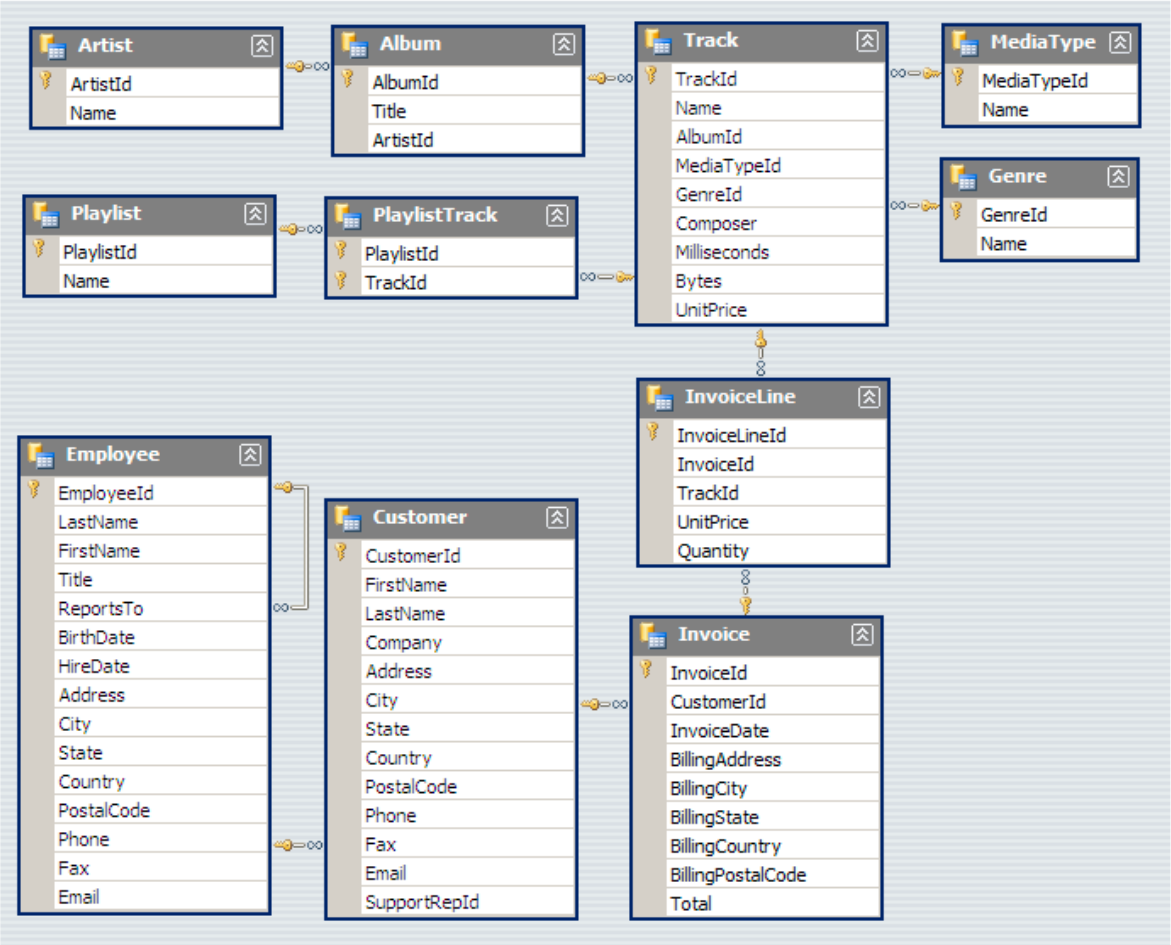
- To understand the sales performance of the music store and identify the top-selling music genres and artists.
- To study that which countries have the most Invoices
- To analyse customer demographics and behaviour to improve customer engagement and loyalty.
- To study the time-based patterns of sales and identify peak hours and days for promotional activities.
- To provide data-driven recommendations to the management for enhancing the music store's performance.
- And many more….

Research Question :

1. Who is the senior most employee based on job title?
2. Which countries have the most Invoices?
3. What are top 3 values of total invoice?
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
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
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
7. 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
8. 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
9. Find how much amount spent by each customer on top artists? Write a query to return customer name, artist name and total spent
10. 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
11. 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

Schema- Music Store Database :



SQL Query :

/* Q1: Who is the senior most employee based on job title? */

SELECT title, last_name, first_name

FROM employee

ORDER BY levels DESC

LIMIT 1

/* Q2: Which countries have the most Invoices? */

SELECT COUNT(*) AS c, billing_country

FROM invoice

GROUP BY billing_country

```
ORDER BY c DESC


/* Q3: What are top 3 values of total invoice? */


SELECT total

FROM invoice

ORDER BY total DESC


/* 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 InvoiceTotal

FROM invoice

GROUP BY billing_city

ORDER BY InvoiceTotal DESC

LIMIT 1;


/* 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.*/


SELECT customer.customer_id, first_name, last_name, SUM(total) AS total_spending

FROM customer

JOIN invoice ON customer.customer_id = invoice.customer_id

GROUP BY customer.customer_id

ORDER BY total_spending DESC

LIMIT 1;


/* 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. */


/*Method 1 */
```

```
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;

/* 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 artist.artist_id, 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;

/* 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) AS avg_track_length
```

```
        FROM track )
ORDER BY milliseconds DESC;


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


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;



/* 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. */


/* Method 1: Using CTE */
```

```
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

ORDER BY purchase DESC


/* 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. */


/* Method 1: using CTE */


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
```

Result :

From the data we got after running the queries. We can say that,

1. The most senior employee based on job title is Madan Mohan.
2. “USA” have the most Invoices.
3. The top 3 values of total invoice are : “23.7599,19.8.19.8”
4. “Prague” has the best customer. And total invoice is 273.24 $
5. “R Madhav” is the best customer and total invoice is 144.54 $
6. I found 59 result who like Genre of all Rock Music listeners. (CSV File Link)
7. The Top 10 Rock music bands name in (CSV file Link)
8. I found 494 result the track names that have a song length longer than the average song length. (CSV File Link)
9. I found 43 customer name who spend on top artist . (CSV File Link)
10. I found top purchase of genre countries wise and I make purchase wise sort decreasing order (CSV File Link)
11. There are 24 countries and I found each county’s top most buying person.(CSV File Link)