# MUSIC STORE DATA ANALYSIS

## 1. Project Overview

The **Music Store Data Analysis Project** focuses on extracting business insights from a relational music store database using SQL. The project analyzes customer purchasing behavior, revenue trends, and music preferences by solving real-world business problems. Various SQL techniques such as joins, subqueries, aggregations, CTEs, and window functions were used to perform multi-table data analysis. The analysis helps identify top customers, best-selling artists, popular genres, and high-revenue locations. This project demonstrates practical data querying and analytical problem-solving skills used in real-world business decision-making.

## 2. Dataset Summary

Rows: ~11 tables (relational schema) – Key Features:

- **Music Catalog Data:** Artist, Album, and Track details including track duration, composer, genre, media type, and pricing.
- **Customer & Employee Information:** Customer demographics, contact details, geographic location, and assigned support representatives.
- **Sales & Transaction Records:** Invoice and InvoiceLine tables capturing purchase transactions, billing details, track purchases, quantity, and total sales amount.
- **Music Classification & Organization:** Genre and MediaType tables categorize tracks, while Playlist and PlaylistTrack tables manage curated music collections.
- **Business Analytics Capability:** Enables analysis of customer purchasing behavior, sales performance, popular music genres, top artists, and revenue trends across countries and cities.

### 3.) Data Analysis Using SQL

To understand the business performance of the music store, the dataset was analyzed using SQL by solving 11 carefully designed business-driven questions grouped into three levels of complexity.

The analysis focused on uncovering valuable insights such as customer purchasing patterns, revenue-generating locations, popular music genres, and top-performing artists. By exploring relationships between customers, sales transactions, and music catalog data, the study highlights how data can support strategic decision-making, improve customer targeting, and optimize sales performance.

This structured approach provides a complete overview of how different components of the business contribute to overall revenue and customer engagement.

**PART 1**

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

```sql
Select
    first_name,
    last_name,
    levels
from employee
ORDER BY Levels DESC
LIMIT 1;

Select *
from genre
```

Data Output | Messages | Notifications

| | first_name character (50) | last_name character (50) | levels character varying (10) |
|---|---|---|---|
| 1 | Mohan | Madan | L7 |

2. Which countries have the most Invoices?

```
Select
    COUNT(*) as c,
    billing_country
from invoice
GROUP BY billing_country
ORDER BY c DESC;
```

Data Output Messages Notificatior

| | c bigint | billing_country character varying (30) |
|---|---|---|
| 1 | 131 | USA |
| 2 | 76 | Canada |
| 3 | 61 | Brazil |
| 4 | 50 | France |
| 5 | 41 | Germany |
| 6 | 30 | Czech Republic |
| 7 | 29 | Portugal |

3. What are top 3 values of total invoice?

```
Select *
from invoice
ORDER BY total DESC
limit 3;
```

Data Output Messages Notifications

Showing rows: 1 to 3 Page No: 1 of 1

| | invoice_id [PK] integer | customer_id integer | invoice_date timestamp without time zone | billing_address character varying (120) | billing_city character varying (30) | billing_state character varying (30) | billing_country character varying (30) | billing_postal character varying (30) |
|---|---|---|---|---|---|---|---|---|
| 1 | 183 | 42 | 2018-02-09 00:00:00 | 9, Place Louis Barthou | Bordeaux | None | France | 33000 |
| 2 | 92 | 32 | 2017-07-02 00:00:00 | 696 Osborne Street | Winnipeg | MB | Canada | R3L 2B9 |
| 3 | 31 | 3 | 2017-02-21 00:00:00 | 1498 rue Bélanger | Montréal | QC | Canada | H2G 1A7 |

Data Output Messages Notifications

Showing rows: 1 to 3 Page No: 1 of 1

| | customer_id integer | invoice_date timestamp without time zone | billing_address character varying (120) | billing_city character varying (30) | billing_state character varying (30) | billing_country character varying (30) | billing_postal character varying (30) | total double precision |
|---|---|---|---|---|---|---|---|---|
| 1 | 42 | 2018-02-09 00:00:00 | 9, Place Louis Barthou | Bordeaux | None | France | 33000 | 23.759999999999998 |
| 2 | 32 | 2017-07-02 00:00:00 | 696 Osborne Street | Winnipeg | MB | Canada | R3L 2B9 | 19.8 |
| 3 | 3 | 2017-02-21 00:00:00 | 1498 rue Bélanger | Montréal | QC | Canada | H2G 1A7 | 19.8 |

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

```
Select
    SUM(total) as invoice_total,
    billing_city
from invoice
GROUP BY billing_city
ORDER BY invoice_total DESC;
```

Data Output Messages Notifications

| | invoice_total double precision | billing_city character varying (30) |
|---|---|---|
| 1 | 273.24000000000007 | Prague |
| 2 | 169.29 | Mountain View |
| 3 | 166.32 | London |
| 4 | 158.4 | Berlin |
| 5 | 151.47 | Paris |
| 6 | 129.69 | São Paulo |
| 7 | 114.83999999999997 | Dublin |

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

```
Select
    c.customer_id,
    c.first_name,
    c.last_name,
    SUM(i.total) as total
from Customer as c
JOIN invoice as i
ON c.customer_id = i.customer_id
GROUP BY c.first_name, c.last_name, c.customer_id
ORDER BY total DESC
LIMIT 1;
```

Data Output Messages Notifications

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

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

```
Select DISTINCT
    c.email,
    c.first_name,
    c.last_name
from Customer as c
JOIN invoice as i
ON c.customer_id = i.customer_id
JOIN invoice_line as il
ON i.invoice_id = il.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 ASC;
```

Data Output Messages Notifications

| | 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 |

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

```
Select
    a.name,
    SUM(t.milliseconds) as total_track_time
from Artist as a
JOIN Album as al
ON a.Artist_Id = al.Artist_Id
JOIN track as t
ON t.Album_Id = al.Album_Id
JOIN genre as g
ON g.genre_Id = t.genre_Id
Where g.Name LIKE 'Rock'
Group by a.name
ORDER BY total_track_time DESC
LIMIT 10;
```

Data Output Messages Notifications

| | name character varying (120) | total_track_time bigint |
|---|---|---|
| 1 | Led Zeppelin | 40121414 |
| 2 | Deep Purple | 32259613 |
| 3 | U2 | 30287097 |
| 4 | Iron Maiden | 30081859 |
| 5 | Pearl Jam | 13518967 |
| 6 | Santana | 12847973 |
| 7 | Van Halen | 12297219 |

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

```
Select
    t.name
from track as t
Where t.milliseconds > (
    Select
    AVG(milliseconds) as avg_track_length
    FROM track
)
ORDER BY milliseconds DESC;
```

Data Output Messages Notifications

Showing ro

| | name character varying (150) |
|---|---|
| 1 | Occupation / Precipice |
| 2 | Through a Looking Glass |
| 3 | Greetings from Earth, Pt. 1 |
| 4 | The Man With Nine Lives |
| 5 | Battlestar Galactica, Pt. 2 |
| 6 | Battlestar Galactica, Pt. 1 |
| 7 | Murder On the Rising Star |

9.

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
        ar.artist_id,
        ar.name AS artist_name,
        SUM(il.unit_price * il.quantity) AS total_sales
    FROM invoice_line il
    JOIN track t
        ON t.track_id = il.track_id
    JOIN album al
        ON al.album_id = t.album_id
    JOIN artist ar
        ON ar.artist_id = al.artist_id
    GROUP BY ar.artist_id, ar.name
    ORDER BY total_sales 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 al
    ON al.album_id = t.album_id
JOIN best_selling_artist bsa
    ON bsa.artist_id = al.artist_id
GROUP BY
    c.customer_id,
    c.first_name,
    c.last_name,
    bsa.artist_name
ORDER BY amount_spent DESC;
```

Data Output Messages Notifications

Showing rows: 1 t

| | customer_id integer | first_name character (50) | last_name character (50) | artist_name character varying (120) | amount_spent 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 |

## 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

```sql
WITH popular_genre AS (
    SELECT
        customer.country,
        genre.name AS genre_name,
        genre.genre_id,
        SUM(invoice_line.quantity) AS purchases,
        ROW_NUMBER() OVER (
            PARTITION BY customer.country
            ORDER BY SUM(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
        customer.country,
        genre.name,
        genre.genre_id
)

SELECT *
FROM popular_genre
WHERE RowNo = 1;
```

Data Output Messages Notifications

| | country character varying (50) | genre_name character varying (120) | genre_id character varying (50) | purchases double precision | rowno bigint |
|---|---|---|---|---|---|
| 1 | Argentina | Alternative & Punk | 4 | 17 | 1 |
| 2 | Australia | Rock | 1 | 34 | 1 |
| 3 | Austria | Rock | 1 | 40 | 1 |
| 4 | Belgium | Rock | 1 | 26 | 1 |
| 5 | Brazil | Rock | 1 | 205 | 1 |
| 6 | Canada | Rock | 1 | 333 | 1 |
| 7 | Chile | Rock | 1 | 61 | 1 |

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

```sql
WITH customer_with_country AS (
    SELECT
        c.customer_id,
        c.first_name,
        c.last_name,
        i.billing_country,
        SUM(i.total) AS total_spending
    FROM invoice i
    JOIN customer c
        ON c.customer_id = i.customer_id
    GROUP BY
        c.customer_id,
        c.first_name,
        c.last_name,
        i.billing_country
),

country_max_spending AS (
    SELECT
        billing_country,
        MAX(total_spending) AS max_spending
    FROM customer_with_country
    GROUP BY billing_country
)
```

```sql
SELECT
    cwc.billing_country,
    cwc.first_name,
    cwc.last_name,
    cwc.total_spending
FROM customer_with_country cwc
JOIN country_max_spending cms
    ON cwc.billing_country = cms.billing_country
   AND cwc.total_spending = cms.max_spending
ORDER BY
    cwc.billing_country;
```

Data Output Messages Notifications

| | billing_country character varying (30) | first_name character (50) | last_name character (50) | total_spending double precision |
|---|---|---|---|---|
| 1 | Argentina | Diego | Gutiérrez | 39.6 |
| 2 | Australia | Mark | Taylor | 81.18 |
| 3 | Austria | Astrid | Gruber | 69.3 |
| 4 | Belgium | Daan | Peeters | 60.38999999999999 |
| 5 | Brazil | Luís | Gonçalves | 108.89999999999998 |
| 6 | Canada | François | Tremblay | 99.99 |
| 7 | Chile | Luis | Rojas | 97.02000000000001 |

## 📌 Business Recommendations & Insights

### 1️⃣ Target High-Revenue Locations with Marketing Campaigns

The analysis identified specific cities and countries generating the highest revenue and invoice counts. The company should prioritize promotional campaigns, concerts, and localized offers in these high-performing regions to maximize revenue and customer engagement.

---

### 2️⃣ Promote Popular Music Genres Based on Regional Preferences

Different countries show varying genre preferences, with certain genres such as Rock attracting a larger audience. The business can improve customer satisfaction and sales by recommending region-specific music genres and tailoring playlists based on geographic listening trends.

---

### 3️⃣ Strengthen Partnerships with Top-Selling Artists

The study highlights artists contributing significantly to overall sales revenue. Collaborating with these artists through exclusive releases, promotions, or featured playlists can help drive higher customer engagement and increase sales.

---

### 4️⃣ Implement Customer Loyalty and Retention Programs

Customer spending analysis revealed that a small group of customers contributes significantly to total revenue. Introducing loyalty programs, personalized discounts, and early access to music releases can help retain high-value customers and increase lifetime value.

---

### 5️⃣ Optimize Product Recommendations Using Purchase Patterns

Customer purchase history and track popularity analysis provide opportunities to implement recommendation systems. Suggesting similar artists, albums, or genres based on customer preferences can improve user experience and boost cross-selling opportunities.

---

### 6️⃣ Utilize Sales Trend Analysis for Strategic Business Planning

Tracking invoice totals and customer spending trends helps identify seasonal or geographic demand patterns. These insights can support inventory planning, pricing strategies, and targeted promotional campaigns to optimize overall sales performance.