# Music Store Data Analysis

***Insights Derived from SQL Analysis***

***Arif Amin***

# CONTENTS

# Introduction

**Project Overview:**

- This project involves analyzing a music store's data using SQL to derive meaningful insights that can enhance business decision-making.

- By performing SQL analysis on this data, we aim to explore key aspects of the store’s performance, including sales trends, customer loyalty, and geographic distribution of sales.

# Data Preparation

**Dataset Overview**:

- Tables analyzed: artist, employee,invoice,customer, etc.
- This dataset comprises of ten tables ; album,artist,customer ,employee, genre, invoice, invoice line, media type, playlist-track and track.

**Steps Taken**:

- Imported data into the SQL database.
- Explored relationships between tables.
- Cleaned and prepared data for queries by removing the special characters and adding null values for blank values.
- Total of 59 customers, 275 artists, 3503 tracks, 347 albums ,9 employees and 614 invoices have been included in the data.

# Objective

The primary goal of this SQL-based data analysis is to uncover insights that can help improve the music store's overall operations, customer engagement, and sales performance by utilizing SQL functions and then visualize the findings using various charts.

**Key Areas of Focus**

- **Customer Analysis:** Understanding customer preferences (genre, spending habits), Identifying top customers, and segmenting customers for targeted marketing.
- **Sales Analysis:** Evaluating sales performance across different regions, identifying top-performing invoices, and analyzing revenue generation.
- **Employee Analysis:** Determining employee seniority.
- **Music Analysis:** Analyzing music genre popularity, artist performance, and track characteristics.

# Methodology

Description of SQL functions used –

- JOIN: To combine data from two or more tables
- DISTINCT: To retrieve unique email ids, first name, last name of customers.
- CTE: For creating temporary result sets.
- Aggregation Functions: SUM, COUNT,AVERAGE, MAX.
- GROUP BY & ORDER BY: For grouping & sorting data.
- WINDOW FUNCTION: ROW NUMBER FUNCTION that assigns a unique sequential integer to each row within a partition of a result set.
- RECURSIVE FUNCTION: The recursive keyword within a CTE where the second query is dependent on first query.

# Employee Analysis

# Who is the senior most employee based on job title.

```
select *
from employee
order by levels desc
limit 1;
```

**INTERPRETATION:**

- **Adams Andrew** is the senior most employee with level 6 (L6).

| employee_id | last_name | first_name | title | reports_to | levels | birthdate | hire_date | address | city | state | country | postal_code | phone | fax | email |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Adams | Andrew | General Manager | 9 | L6 | 18-02-1962 00:00 | 14-08-2016 00:00 | 11120 Jasper Ave NW | Edmonton | AB | Canada | T5K 2N1 | +1 (780) 428-9482 | +1 (780) 428-3457 | andrew@chinookcorp.com |

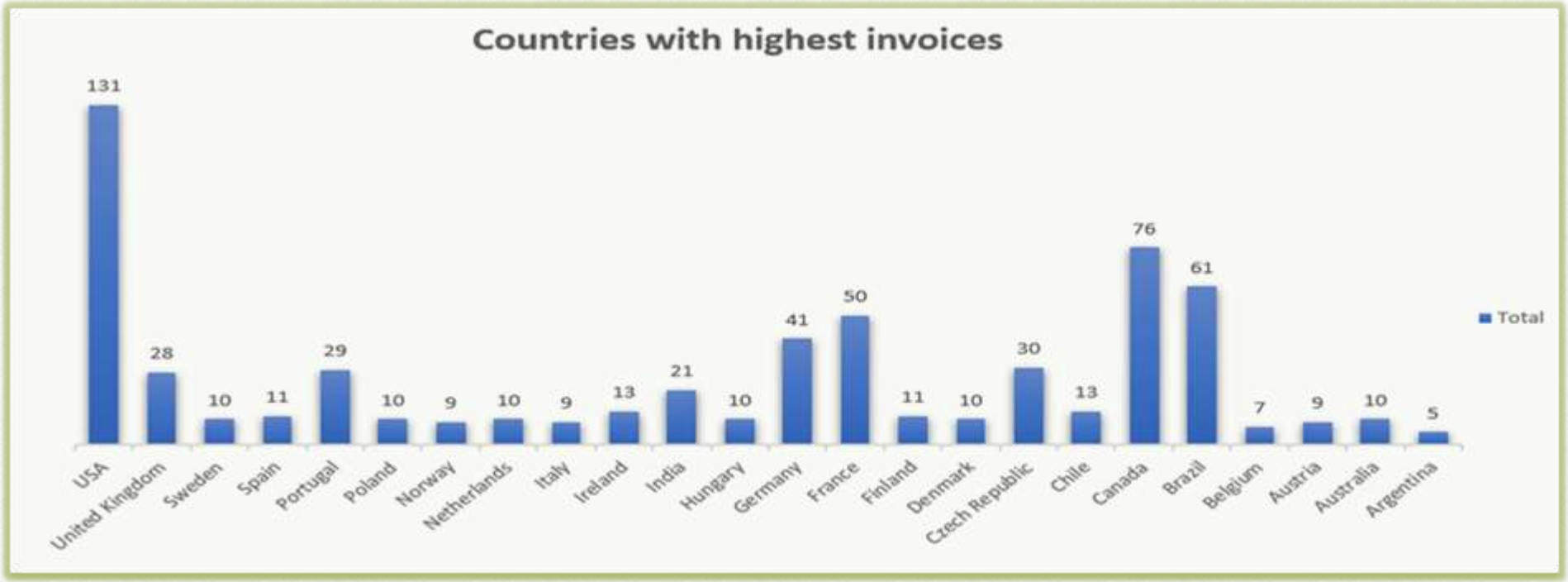
# Sales Analysis

# Which Countries have the most invoices

```
select billing_country,count(*) as c
from invoice
group by billing_country
order by c desc;
```

**INTERPRETATION:**

**USA** has the highest invoice total of **131.**

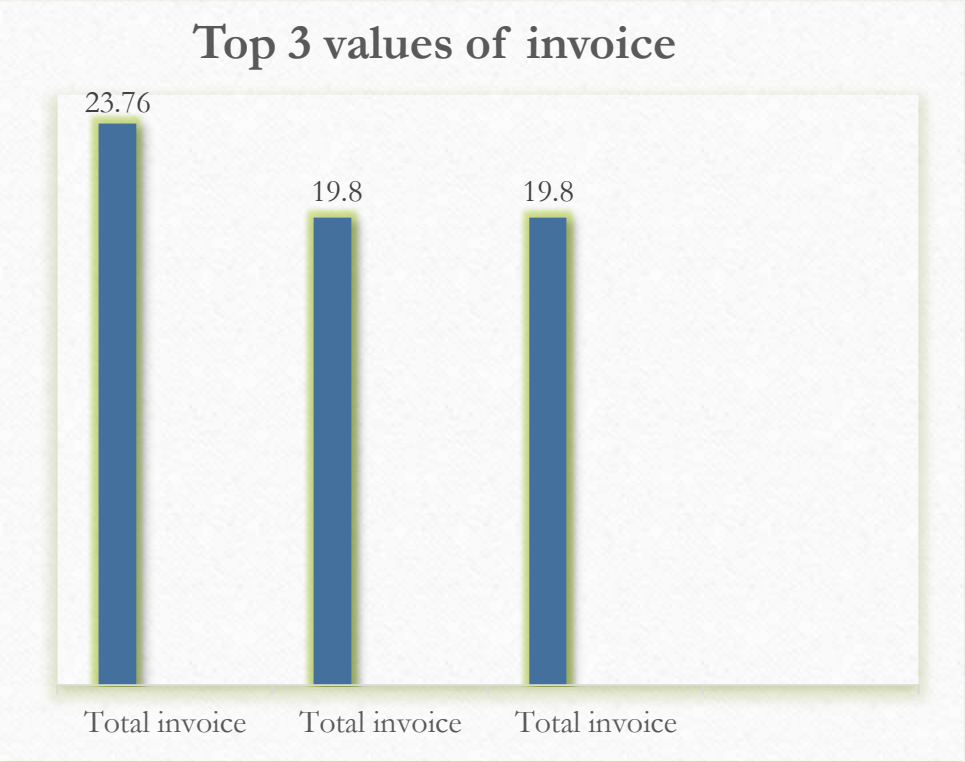


# What are top 3 values of total invoice

```
select total
from invoice
order by total desc
limit 3;
```

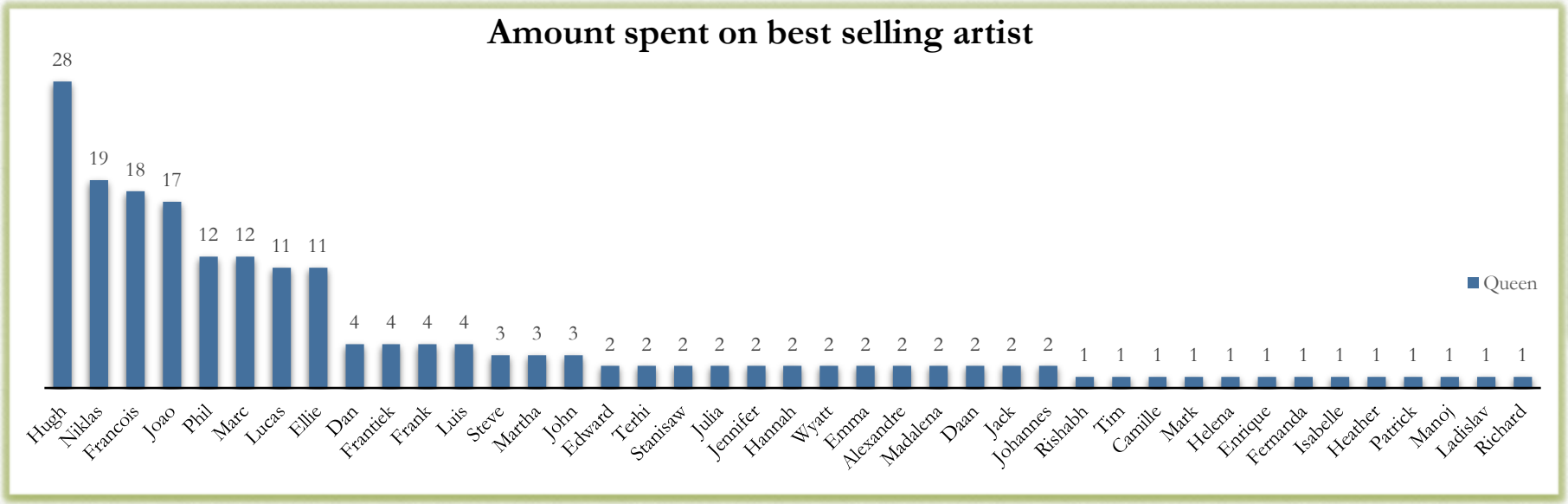
**INTERPRETATION:**

**The top invoice total value is 23.76, followed by 19.8 twice.**



# Amount spent by each customer on best selling artists

```sql
With best_selling_artist AS(
select a.artist_id,a.name,sum(il.unit_price*il.quantity) as total_sales
from invoice_line il join track t on il.track_id=t.track_id join album al on al.album_id=t.album_id join artist a on a.artist_id=al.artist_id
group by 1,2
order by 3 desc
limit 1)
select c.customer_id,c.first_name,c.last_name,bsa.name,sum(il.unit_price*il.quantity) as amount_spent
from customer c join invoice i 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;
```




**INTERPRETATION:**

**Hugh** has spent the highest amount '28'on the best selling artist "Queen".

# Customer Analysis

# 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 money
from customer c  inner join invoice i on c.customer_id=i.customer_id
group by c.customer_id,c.first_name,c.last_name
order by money desc
limit 1;
```

| customer_id | first_name | last_name | money |
|---|---|---|---|
| 5 | Frantiek | Wichterlova | 144.54000000000002 |

**INTERPRETATION:**
The customer who has spent the most money is Franteik Wichterlova

# Write a 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 email,first_name, last_name
from customer c join invoice i on c.customer_id=i.invoice_id
join invoice_line il on il.invoice_id=i.invoice_id
where track_id in(
Select t.track_id
from track t join genre g on t.genre_id=g.genre_id
WHERE g.name LIKE '%Rock%'
)
order by email;
```

| email | first_name | last_name |
|---|---|---|
| aaronmitchell@yahoo.ca | Aaron | Mitchell |
| alero@uol.com.br | Alexandre | Rocha |
| astrid.gruber@apple.at | Astrid | Gruber |
| bjorn.hansen@yahoo.no | Bjrn | Hansen |
| camille.bernard@yahoo.fr | Camille | Bernard |
| daan_peeters@apple.be | Daan | Peeters |
| diego.gutierrez@yahoo.ar | Diego | Gutierrez |
| dmiller@comcast.com | Dan | Miller |
| dominiquelefebvre@gmail.com | Dominique | Lefebvre |
| edfrancis@yachoo.ca | Edward | Francis |
| eduardo@woodstock.com.br | Eduardo | Martins |
| ellie.sullivan@shaw.ca | Ellie | Sullivan |
| emma_jones@hotmail.com | Emma | Jones |
| enrique_munoz@yahoo.es | Enrique | Munoz |
| fernadaramos4@uol.com.br | Fernanda | Ramos |
| fharris@google.com | Frank | Harris |
| fralston@gmail.com | Frank | Ralston |
| frantisekw@jetbrains.com | Frantiek | Wichterlova |
| ftremblay@gmail.com | Francois | Tremblay |
| fzimmermann@yahoo.de | Fynn | Zimmermann |
| hannah.schneider@yahoo.de | Hannah | Schneider |
| hholy@gmail.com | Helena | Hol |
| hleacock@gmail.com | Heather | Leacock |
| hughoreilly@apple.ie | Hugh | O'Reilly |
| isabelle_mercier@apple.fr | Isabelle | Mercier |
| jacksmith@microsoft.com | Jack | Smith |
| jenniferp@rogers.ca | Jennifer | Peterson |

**INTERPRETATION:**

Customers who are rock music listeners are '55' in number.

# Write a query that returns the country along with the top customer and how much they spent.

```
with customer_with_country as(
select c.customer_id,c.first_name,c.last_name,i.billing_country,sum(i.total) as total_spending,
row_number() over(partition by billing_country order by sum(total) desc) as row_no
from customer c join invoice i on c.customer_id=i.customer_id
group by 1,2,3,4
order by 4 asc,5 desc)
select * from customer_with_country
where row_no <=1;
```

| customer_id | first_name | last_name | billing_country | total_spending | row_no |
|---|---|---|---|---|---|
| 56 | Diego | Gutierrez | Argentina | 39.6 | 1 |
| 55 | Mark | Taylor | Australia | 81.18 | 1 |
| 7 | Astrid | Gruber | Austria | 69.3 | 1 |
| 8 | Daan | Peeters | Belgium | 60.39 | 1 |
| 1 | Luis | Goncalves | Bra Brazil | 108.89999999999999 | 1 |
| 3 | Francois | Tremblay | Canada | 99.99 | 1 |
| 57 | Luis | Rojas | Chile | 97.02000000000001 | 1 |
| 5 | Frantiek | Wichterlova | Czech Republic | 144.54000000000002 | 1 |
| 9 | Kara | Nielsen | Denmark | 37.61999999999999 | 1 |
| 44 | Terhi | Hmlinen | Finland | 79.2 | 1 |
| 42 | Wyatt | Girard | France | 99.99 | 1 |
| 37 | Fynn | Zimmermann | Germany | 94.05000000000001 | 1 |
| 45 | Ladislav | Kovacs | Hungary | 78.21 | 1 |
| 58 | Manoj | Pareek | India | 111.86999999999999 | 1 |
| 46 | Hugh | O'Reilly | Ireland | 114.83999999999997 | 1 |
| 47 | Lucas | Mancini | Italy | 50.49 | 1 |
| 48 | Johannes | Van der Berg | Netherlands | 65.34 | 1 |
| 4 | Bjrn | Hansen | Norway | 72.27000000000001 | 1 |
| 49 | Stanisaw | Wjcik | Poland | 76.22999999999999 | 1 |
| 34 | Joao | Fernandes | Portugal | 102.96000000000001 | 1 |
| 50 | Enrique | Munoz | Spain | 98.01 | 1 |
| 51 | Joakim | Johansson | Sweden | 75.24 | 1 |
| 53 | Phil | Hughes | United Kingdom | 98.01 | 1 |
| 17 | Jack | Smith | USA | 98.01 | 1 |

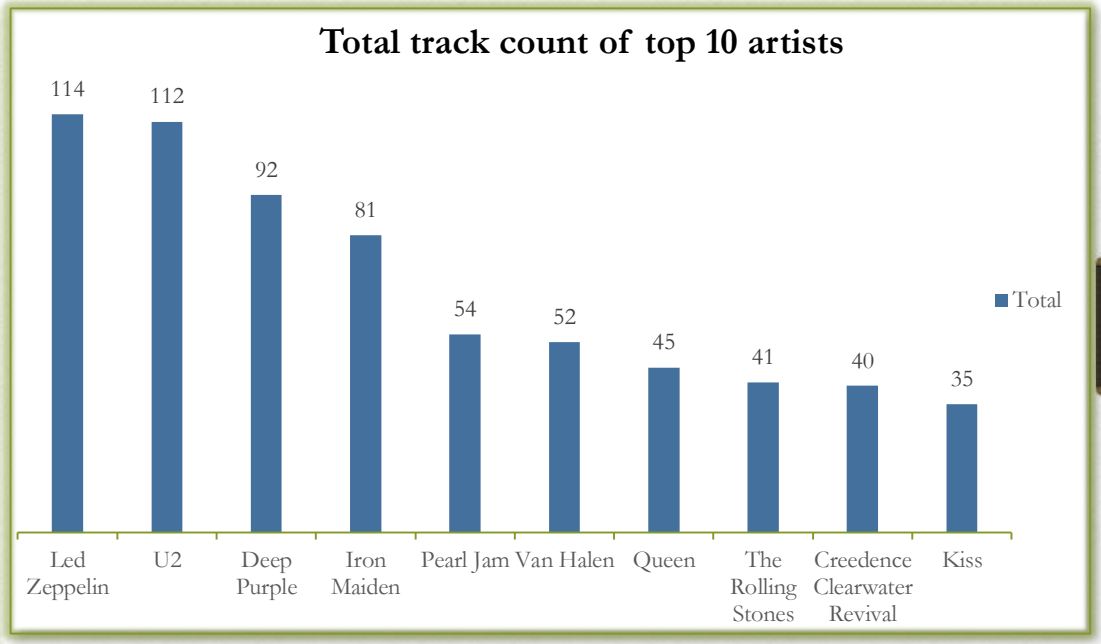
**INTERPRETATION:**

Gives the top customers for each country which are "24" in number

# Music Analysis

## Write a query that returns the Artist name and total track count of the top 10 rock bands.

```sql
Select a.artist_id,a.name,count(t.track_id) as Total_track_count
from track t join album al on al.album_id=t.album_id
join artist a on a.artist_id=al.artist_id
where track_id in(
 Select t.track_id
 from track t join genre g on t.genre_id=g.genre_id
WHERE g.name LIKE '%Rock%'
)
group by a.artist_id,a.name
order by Total_track_count desc
limit 10;
```




**INTERPRETATION:**

Led Zeppelin is the artist with highest track count from Rock genre

**Return all the track names that have a song length longer than the average song length.**

```
select name,milliseconds as song_length
from track
where milliseconds >(
select avg(milliseconds) as avg_song_length
from track)
order by milliseconds desc;
```

**INTERPRETATION:**

The track names with length greater than average song length(393599.2121) are returned which are "424" in number.

| name | song_length |
|---|---|
| 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 |
| Murder On the Rising Star | 2935894 |
| Battlestar Galactica Pt 3 | 2927802 |
| Take the Celestra | 2927677 |
| Fire In Space | 2926593 |
| The Long Patrol | 2925008 |
| The Magnificent Warriors | 2924716 |
| The Living Legend Pt 1 | 2924507 |
| The Gun On Ice Planet Z... | 2924341 |
| The Hand of God | 2924007 |
| Experiment In Terra | 2923548 |
| War of the Gods Pt 2 | 2923381 |
| The Living Legend Pt 2 | 2923298 |
| War of the Gods Pt 1 | 2922630 |
| Lost Planet of the Gods P... | 2922547 |
| Baltars Escape | 2922088 |
| The Lost Warrior | 2920045 |
| Lost Planet of the Gods P... | 2914664 |
| The Gun On Ice Planet Z... | 2907615 |
| Greetings from Earth Pt 2 | 2903778 |

# Challenges and Learnings

- Challenges faced during analysis:
  - Understanding complex relationships between tables.
  - Cleaning and preparing the data.
- Key Learnings:
  - How SQL queries can extract valuable insights.
  - Importance of exploring and visualizing data.

# Recommendations/Insightful suggestions

- **Customer Engagement and Personalization**

  **Top Spenders & Loyal Customers:** Since Frantiek Wichterlova has been identified as the top customer, the store should create personalized loyalty programs or offer exclusive discounts to such top spenders.

  **Targeted Genre Promotions:** Many customers prefer Rock music. The store could promote rock albums, host rock- themed events, or offer bundled deals focused on rock music to increase sales in this genre.

- **Geographic Marketing Strategy**

  **Focus on High-Invoice Countries:** Since the USA has the highest number of invoices, the store can concentrate marketing and promotional campaigns there to maximize Return on Investment (ROI). This could include regional events, social media campaigns, or localized offers.

  **Expand to Underperforming Regions:** Conduct further analysis to identify underperforming regions or cities and explore ways to attract new customers there, possibly through partnerships with local artists or record stores.

# Recommendations/Insightful suggestions

- **Artist and Track-Based Insights**

  **Leverage Popular Artists:** Queen has been identified as a best-selling artist. The store could create promotional campaigns, such as "Artist of the Month" featuring Queen's albums, exclusive merchandise, or themed playlists to drive further sales.

  **Highlight Long Tracks:** Promote tracks with lengths greater than the average song length, as they may provide more value to customers who enjoy extended listening experiences.

- **Improving Employee Effectiveness**

  **Employee Recognition:** Highlighting "Adam Andrew "as the senior-most employee (L6) suggests a structured hierarchy. Recognizing and rewarding experienced employees like Adam can boost morale and productivity.

  **Training Programs:** Use the experience of senior employees to mentor junior staff and enhance customer interactions, sales techniques, and product knowledge.

# Recommendations/Insightful suggestions

- **Sales Performance Optimization**

  **Analyze Invoice Trends:** Identify the reasons behind the top three invoice values (23.76, 19.8, 19.8) and explore if similar patterns (product bundles, special deals) can be leveraged to encourage customers to make higher-value purchases.

- By implementing these strategies, the music store can improve its customer retention, expand its geographic reach, and boost overall revenue.

# THANK YOU