MUSIC STORE ANALYSIS

MYSQL PROJECT
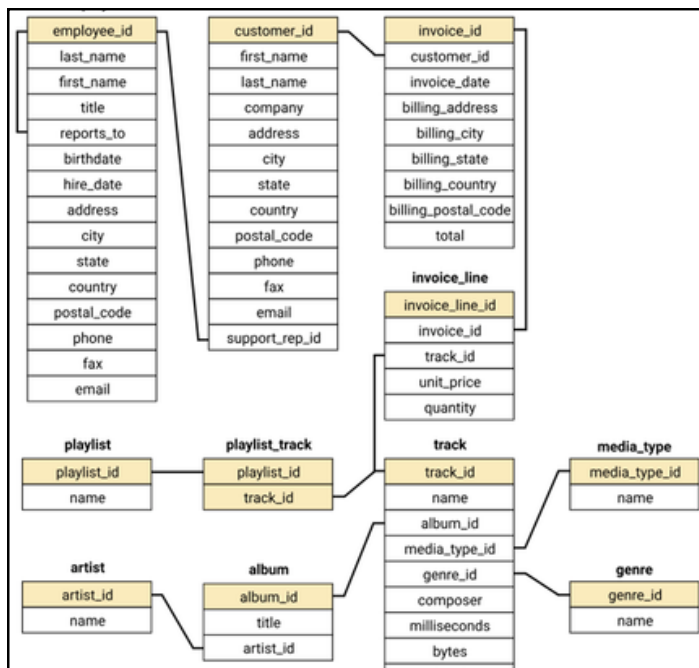
Introduction

- The store has provided data across several tables, including "album.csv", "album2.csv", "artist.csv", "customer.csv", and "playlist.csv" etc. Each dataset contains valuable information that can be leveraged to gain insights into different facets of the business.

Problem Statement

- The objective of this project is to help the Music Store understand its business growth by answering key questions.

Music database

By Mou Sarkar

MUSIC STORE ANALYSIS

q1) Who is the senior most employee based on job title?

- Input:

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

- Output

employee_id	last_name	first_name	title	reports_to	levels	birthdate	hire_date
8	Madan	Mohan	Senior General Manager	NULL	L7	26-01-1961 00:00	14-01-2016 00:00

q2) Which countries have the most Invoices?

- Input:

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

- Output

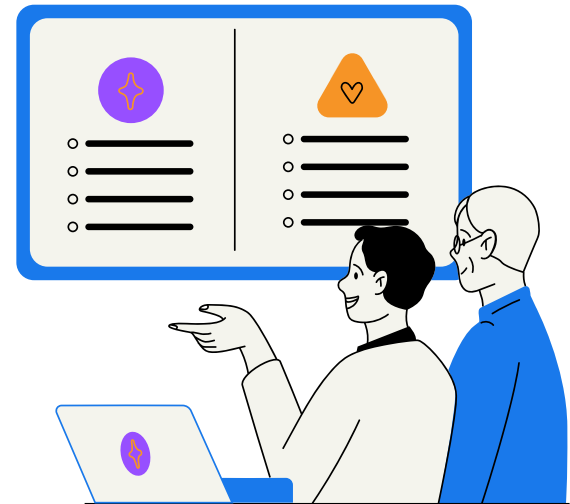| total	billing_country
131	USA
76	Canada
61	Brazil
50	France
41	Germany
30	Czech Republic
29	Portugal
28	United Kingdom
21	India
13	Ireland

By Mou Sarkar

MUSIC STORE ANALYSIS

Q3) What are top 3 values of total invoice?

- Input:

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

- Output

total
23.76
19.8
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.

- Input:

```
select billing_city, sum(total) total_invoice from invoice
group by  billing_city
order by total_invoice desc limit 1;
```

- Output

billing_city	total_invoice
Prague	273.24000000000007

By Mou Sarkar

MUSIC STORE ANALYSIS

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.

- Input:

```
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.customer_id,c.first_name,c.last_name
order by total desc limit 1;
```

- Output

customer_id	first_name	last_name	total
5	František	Wichterlová	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.

- Input:

```
select distinct c.email,c.first_name,c.last_name,g.name
from customer as c
join invoice as i on i.customer_id= c.customer_id
join invoice_line on i.invoice_id= invoice_line.invoice_id
join track as t on t.track_id= invoice_line.track_id
join genre as g on g.genre_id= t.genre_id
where g.name like 'Rock'
order by c.email asc;
```

MUSIC STORE ANALYSIS

- Output

email	first_name	last_name	name
aaronmitchell@yahoo.ca	Aaron	Mitchell	Rock
alero@uol.com.br	Alexandre	Rocha	Rock
astrid.gruber@apple.at	Astrid	Gruber	Rock
bjorn.hansen@yahoo.no	Bjørn	Hansen	Rock
camille.bernard@yahoo.fr	Camille	Bernard	Rock
daan_peeters@apple.be	Daan	Peeters	Rock
diego.gutierrez@yahoo.ar	Diego	Gutiérrez	Rock
dmiller@comcast.com	Dan	Miller	Rock
dominiquelefebvre@gmail.com	Dominique	Lefebvre	Rock

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.

- Input:

```
select ar.artist_id,ar.name,count(ar.artist_id)
as num_of_songs from track t
join album al on al.album_id= t.album_id
join artist ar on ar.artist_id= al.artist_id
join genre g on g.genre_id= t.genre_id
where g.name like 'Rock'
group by ar.artist_id, ar.name
order by num_of_songs desc limit 10;
```

- Output

artist_id	name	num_of_songs
22	Led Zeppelin	114
150	U2	112
58	Deep Purple	92
90	Iron Maiden	81
118	Pearl Jam	54
152	Van Halen	52
51	Queen	45
142	The Rolling Stones	41
76	Creedence Clearwater Revival	40
52	Kiss	35

MUSIC STORE ANALYSIS

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.

- Input:

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

- Output

name	milliseconds
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

MUSIC STORE ANALYSIS

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

- Input:

```
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 artist.artist_id, artist.name
  ORDER BY total_sales DESC
  LIMIT 1
)
select customer.customer_id,customer.first_name,customer.last_name,
bsa.artist_name,
SUM(invoice_line.unit_price * invoice_line.quantity) AS total_amount_spent
from invoice
join customer on customer.customer_id= invoice.customer_id
join invoice_line on invoice_line.invoice_id = invoice.invoice_id
join track on track.track_id= invoice_line.track_id
join album on album.album_id = track.album_id
join best_selling_artist bsa on bsa.artist_id = album.artist_id
group by customer.customer_id,customer.first_name,customer.last_name,
bsa.artist_name
order by  total_amount_spent desc;
```

MUSIC STORE ANALYSIS

- Output

customer_id	first_name	last_name	artist_name	total_amount_spent
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
41	Marc	Dubois	Queen	11.88
47	Lucas	Mancini	Queen	10.89
33	Ellie	Sullivan	Queen	10.89
20	Dan	Miller	Queen	3.96
5	František	Wichterlová	Queen	3.96

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.

- Input:

```
with popular_genre as
(select count(il.quantity) as purcheses,c.country, g.name, g.genre_id,
row_number() over(partition by c.country
order by count(il.quantity) desc) as rowno from invoice_line il
join invoice i on i.invoice_id= il.invoice_id
join customer c on c.customer_id= i.customer_id
join track t on t.track_id= il.track_id
join genre g on g.genre_id= t.genre_id
group by 2,3,4
order by 2 asc, 1 desc)
select * from popular_genre
where rowno <=1;
```

MUSIC STORE ANALYSIS

purcheses	country	name	genre_id	rowno
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
333	Canada	Rock	1	1
61	Chile	Rock	1	1
143	Czech Republic	Rock	1	1
24	Denmark	Rock	1	1
46	Finland	Rock	1	1

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.

- Input:

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

MUSIC STORE ANALYSIS

- Output

customer_id	first_name	last_name	billing_country	total_spending	rowno
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	60.39	1
1	Luís	Gonçalves	Brazil	108.89999999999999	1
3	François	Tremblay	Canada	99.99	1
57	Luis	Rojas	Chile	97.02000000000001	1
5	František	Wichterlová	Czech Republic	144.54000000000002	1
9	Kara	Nielsen	Denmark	37.61999999999999	1
44	Terhi	Hämäläinen	Finland	79.2	1

Results:

- Madan Mohan is the most senior employee based on job title.
- The top country for music purchases is the USA.
- Prague has the best customers, with a total invoice count of around 274.
- The best customer is Frantisek Wichterlova, who has spent approximately $145.
- Led Zeppelin and U2 have written the most rock music.
- The longest song in the playlist is 'Occupation / Precipice.'
- The most popular genre is Rock.

THANK YOU