# Answering Business Questions using SQL

This project aims to analyze and answer key business questions about a fictional music record store called klook Music Players Inc ( KMP)

## Data set

**Database :: music store database**

Please download the .sql file from the link mentioned here and restore the database into your postgreSQL editor / pgAdmin

https://drive.google.com/drive/1J9XBNPqE2FmEjzm5iryQU-EJPQLDrroZ?export=download

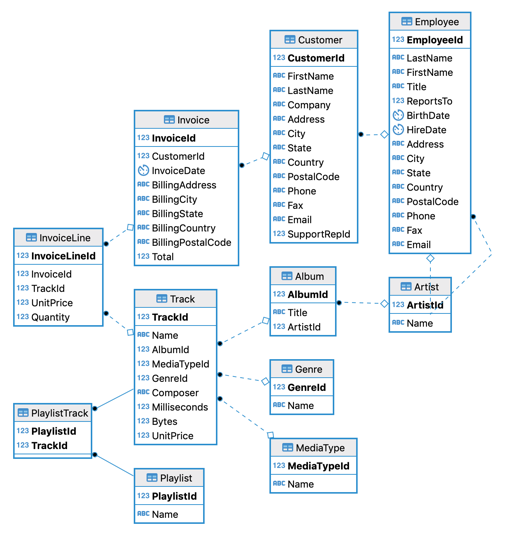
This database includes tables on invoice information, track, album, artist and genre data, and employee and customer information related to the a music store's sales

There are 11 tables in the music store database.

* **employee** table stores employees data such as employee id, last name, first name, etc. It also has a field named **ReportsTo** to specify who reports to whom.
* **customer** table has data related to customers, customers name company , email id, phone their service representative for any query resolution etc.
* **invoice & invoiceLine** tables: these two tables store invoice data.The invoices table stores invoice header data and the invoice\_items table stores the invoice line items data. You can check on total sales or billing related analysis here
* **artist** table stores artists data. It is a simple table that contains only the artist id and name of the music artists.
* **album** table stores data about a list of tracks. Each album belongs to one artist. However, one artist may have multiple albums.
* **media\_type** table stores media types such as MPEG audio and AAC audio files.
* **genre** table stores music types such as rock, jazz, metal,classical etc.
* **track** table stores the data of songs, each of the tracks belongs to one album.
* **playlist & playlist\_track** tables: playlists table store data about playlists. Each playlist contains a list of tracks. Each track may belong to multiple playlists. The relationship between the playlists table and tracks table is many-to-many. The playlist\_track table is used to reflect this relationship.

## Schema

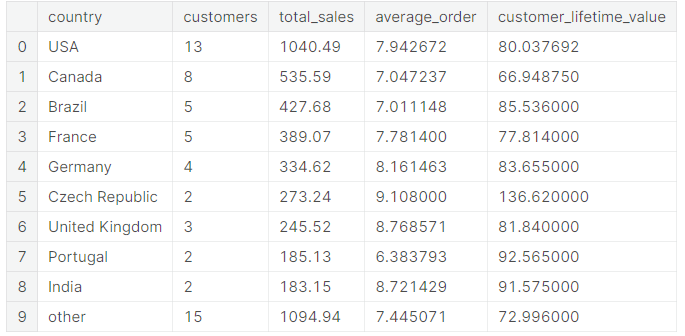
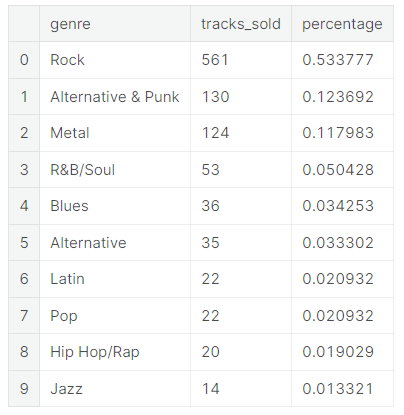
You can refer the below schema to understand the various relationships and keys



## Concepts Covered

CONCAT, WITH,SUBQUERY, SELF JOIN, AVERAGE, AGGREGATION GROUP BY,CASE WHEN , CASTING , HAVING, INNER AND LEFT JOINS

## Business Questions

1. Order of employees by date they joined? Name the person who joined recently **[Beginners]**
2. Who is the oldest employee in the firm? **[Beginners]**
3. What is the average duration of songs in minutes? **[Beginners]**
4. Who is the customer that has spent the most on music for each country? **[Intermediate]**
5. Which artist has the longest songs? **[Beginners]**
6. Find the total dollar amount of sales assigned to each sales support agent within the company **[Intermediate]**
7. Write a SQL query to prepare the following summary - ( country,customers count, total sales amount, average order value and customer lifetime value **[Pro]**
   1. 
   2. Customer lifetime value CLTV : Customer lifetime value is the total worth to a business of a customer over the whole period of their relationship, here you can use total sales per customers in their whole lifespan of buying tracks from the KLOOK store
8. Which genres sell the most tracks in the USA, return the following table- Top 10 genres, tracks sold and their percentage contribution in the market.**[Intermediate]**
   1. 

## Evaluation Rubric

| **Rubric Description** | **Poor** | **Fair** | **Good** |
| --- | --- | --- | --- |
| **How Correct is the Output** | Queries do not work correctly, output is wrong or there is no output | Query structured correctly - logically but there is few small error which were not handled | All queries work correctly |
| **SQL Query Structure and logic building** | Most queries are in long lines, new lines aren't used properly (e.g. SELECT queries are not broken by SELECT, FROM, WHERE, etc.) | Query structure needs work | Query structured logically aligned with Business Knowledge |
| **How closely is the naming convention are followed** | Names are too short or too cryptic, inconsistent | Names follow the conventions in most part, but need work | Alias / names are easy to understand, consistent |
| **Correctness of Joins, Filters and Conditional Statements** | Incorrect usage of concepts | Joins,Filters are correct but need some more work | All joins, filter and conditions are placed correctly |
| **Using Advanced SQL Concepts**  **( With, Sub query)** | With, Sub query are not used and output is not coming correct | With, Sub query are not used but output correct  Need to improve on this to make query more efficient | With, Sub query are used properly and correct output |

## Marks Matrix

For all correct responses maximum marks is as given below

| **Level** | **Max Marks** | **Poor Contr of Max Marks** | **Fair contr of Max Marks** |
| --- | --- | --- | --- |
| Beeginers | 10 | 20% | 60% |
| Intermediate | 15 | 20% | 60% |
| Pro | 25 | 20% | 60% |

## Answers & Explanations ()

1. Order of employees by date they joined
   1. select (last\_name || ' '|| first\_name)as employee\_name,

date(hire\_date) as hiredate

from employee

order by hire\_date asc

1. Who is the oldest employee in the firm?
   1. select (last\_name || ' '|| first\_name)as employee\_name,date(birthdate) as birthdate

from employee

order by birthdate asc

1. What is the average duration of songs in minutes?
   1. select round(avg(milliseconds)/60000,2) as Average\_Duration

from track limit 3

1. Who is the customer that has spent the most on music for each country?
   1. with t1 as

(

select cu.Country as Countr, cu.Customer\_Id as \_Id

, cu.first\_name as F\_name, cu.Last\_Name as L\_name,

sum(i.total) as tm

from Invoice i

JOIN customer cu

on cu.Customer\_Id = i.Customer\_Id

group by 1,2 , 3,4

order by 5),

t2 as

(select countr , max(tm) as mm from t1

group by 1 )

select t1.countr as Country , t1.tm as Totalexpense ,

t1.f\_name || ' ' || t1.L\_name as Full\_name , t1.\_Id as Customer\_Id

from t1

join t2

on t1.countr = t2.countr and t1.tm = t2.mm

order by 2 desc

1. Which artist has the longest songs
   1. select composer, round(avg(milliseconds)/60000,2) as

Average\_Duration

from track

where composer is not null

group by composer

having count(name)>1

order by Average\_Duration desc limit 3

1. Analyzing Employee Sales Performance

Each customer for the store gets assigned to a sales support agent within the company when they first make a purchase. analyze the purchases of customers belonging to each employee to see if any sales support agent is performing either better or worse than the others

find the total dollar amount of sales assigned to each sales support agent within the company

* 1. Select e.first\_name || '' || e.last\_name employee, e.hire\_date,

SUM(inv.total) total\_sales from employee e

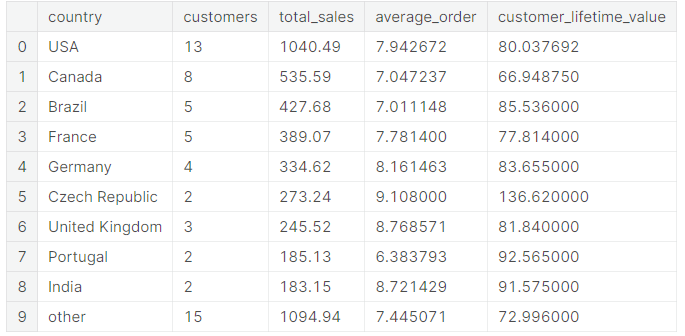
left join customer c

ON c.support\_rep\_id = e.employee\_id

Inner Join invoice inv

On inv.customer\_id = c.customer\_id::int

Group by 1,2

1. calculate data, for each country, on the:
   1. total number of customers
   2. total value of sales
   3. average value of sales per customer
   4. average order value
   5. Expected Output
   6. 
   7. Answer
   8. With purchase\_data As (

Select

c.country,

COUNT(Distinct(inv.customer\_id)) customers,

SUM(inv.total) total\_sales,

SUM(inv.total)/COUNT(inv.customer\_id) average\_order,

SUM(inv.total)/COUNT(Distinct(inv.customer\_id)) customer\_lifetime\_value

From customer c

Inner Join invoice inv

ON c.customer\_id = inv.customer\_id

Group BY 1

Order By 3 Desc),

purchases As (

Select

CASE

when customers = 1 Then 'other'

Else country

End As country,

customers,

total\_sales,

average\_order,

customer\_lifetime\_value

from purchase\_data),

updated As(

Select country,Sum(customers) customers, Sum(total\_sales) total\_sales,

Sum(average\_order)/Count(average\_order) average\_order,

Sum(customer\_lifetime\_value)/count(customer\_lifetime\_value) customer\_lifetime\_value

from purchases

Group by 1

Order BY 3 Desc),

final as (Select \*,

Case When country = 'other' Then 1

Else 0

End as ord

From updated order by ord)

Select country,customers,total\_sales,average\_order,

customer\_lifetime\_value

From final

1. Which genres sell the most tracks in the USA, return the table given in the question- Top 10 genres, tracks sold and their percentage contribution in the market

WITH

genre\_track\_sold AS

(

SELECT

g.name genre,

il.quantity,

il.invoice\_id

FROM genre g

INNER JOIN track t ON g.genre\_id = t.genre\_id

INNER JOIN invoice\_line il ON t.track\_id = il.track\_id

),

sold\_USA AS

(

SELECT

gts.genre,

gts.quantity,

c.country

FROM genre\_track\_sold gts

INNER JOIN invoice i ON i.invoice\_id = gts.invoice\_id

INNER JOIN customer c ON c.customer\_id = i.customer\_id

WHERE country = 'USA'

)

SELECT

genre,

SUM(quantity) tracks\_sold,

CAST(SUM(quantity) as float) / (SELECT COUNT(\*) FROM sold\_USA) percentage

FROM sold\_USA

GROUP BY 1

ORDER BY 2 DESC

LIMIT 10;