**Case Study**

**SQL Project**

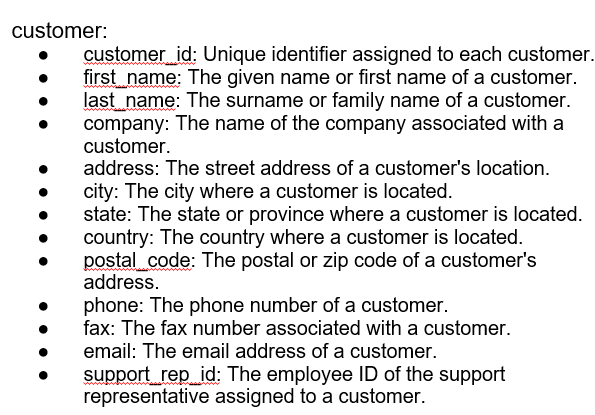
**Chinook Music Store**

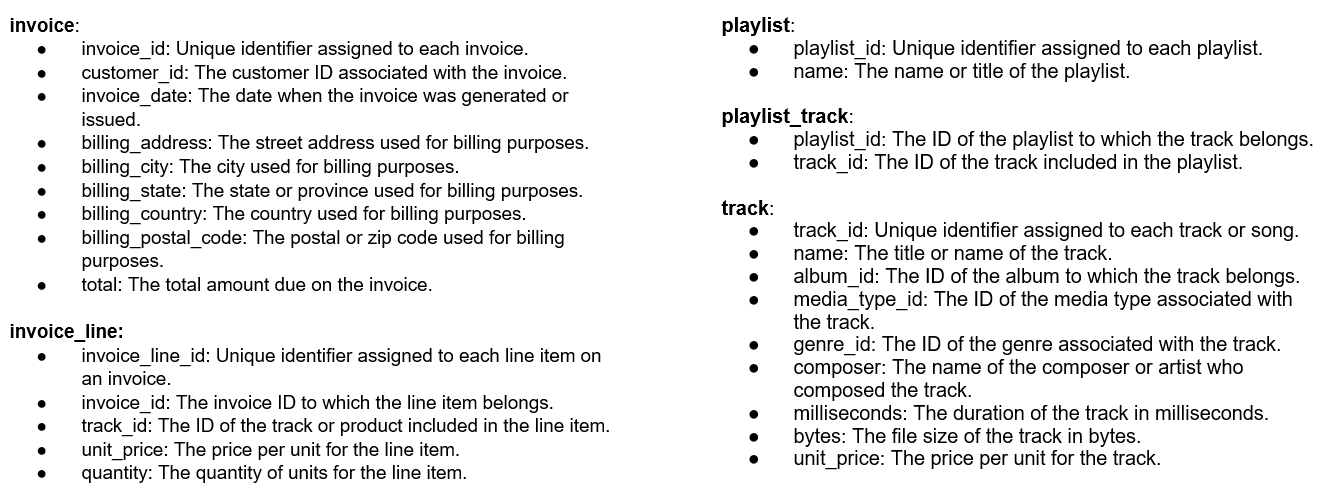
**Project by:** Tammana Sai Manikanta **Newton Batch:** April 2025 **Date:** 20-July-2025

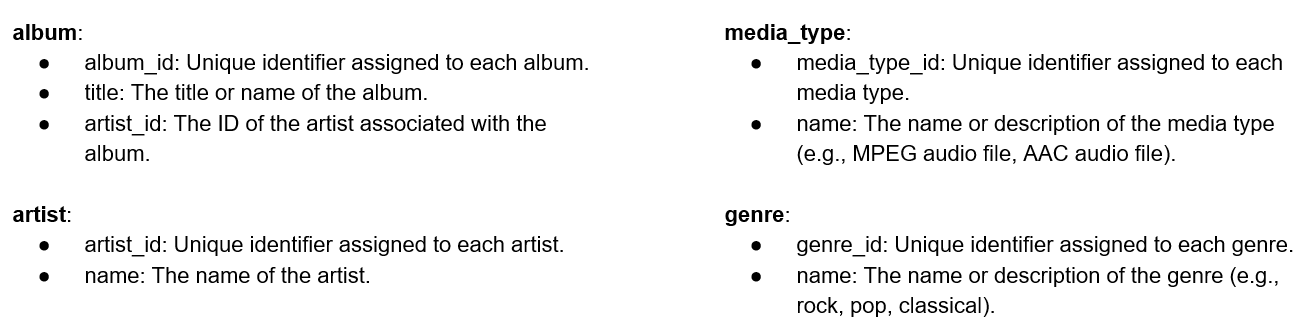
**Tools Used:** MySQL Workbench

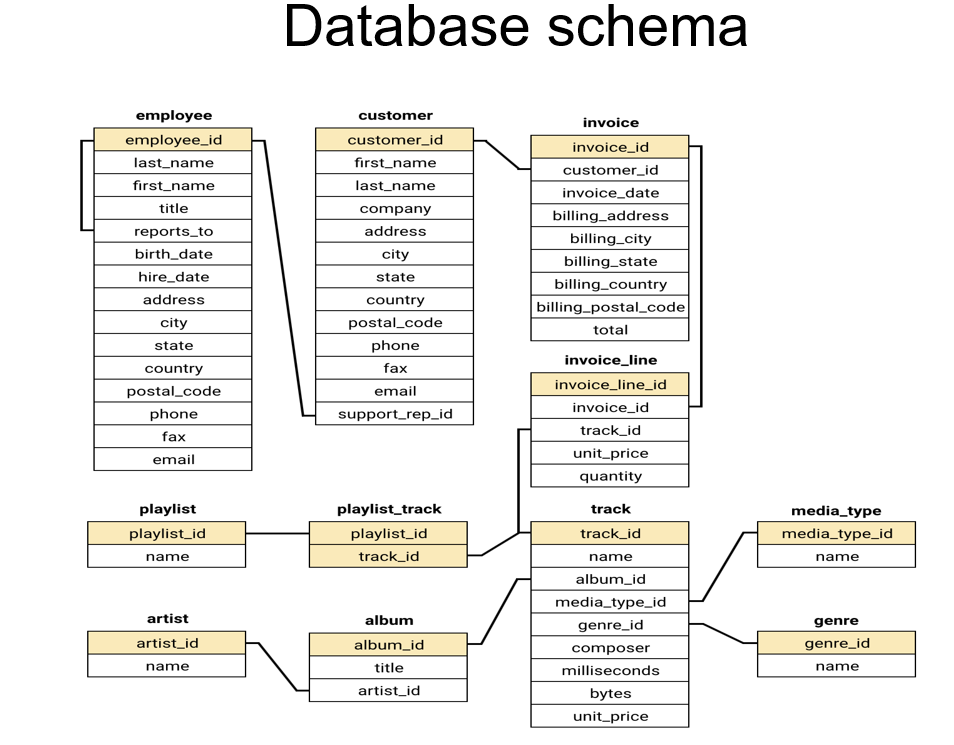
**Learners have to come up with a Report to support the answers to the following questions and suggestions**

Database Consists of









Objective Questions

1. Does any table have missing values or duplicates? If yes, how would you handle it ?

use chinook;

select \* from album;

select \* from artist;

select \* from customer;

select \* from employee;

select \* from genre;

select \* from invoice;

select \* from invoice\_line;

select \* from media\_type;

select \* from playlist;

select \* from playlist\_track;

select \* from track;

I have exported every table in csv to check the Null values exists or not and then I have replaced those Null values using Coalesce function.

* Created a view for the tables where the null values are existing and to recall for future reference as we can consider the NA for Null values

For these tables mentioned below

album, artist, employee, genre, invoice, invoice\_line, media\_type, playlist, playlist\_track we do not have any null values

-employee table we have one null value where employee1 is reporting to null as he may be the chairman or CEO for the organization

customer:

create view customer\_view as

select customer\_id, first\_name, last\_name, coalesce(company, "NA") as company,

coalesce(address, "NA") as address, city, coalesce(state, "NA") as state,

country, coalesce(postal\_code,"NA") as postal\_code, coalesce(phone,"NA") as phone, coalesce(fax,"NA") as fax, email, support\_rep\_id from customer;

track:

create view track\_view

as select track\_id, name, album\_id, media\_type\_id, genre\_id, coalesce(composer,"NA") as composer,

milliseconds, bytes, unit\_price

from track;

1. Find the top-selling tracks and top artist in the USA and identify their most famous genres.

--Top-selling tracks in USA

select t.track\_id, t.name as track\_name, sum(il.quantity) as total\_sold, g.name as genre\_name, a.name as artist from

invoice\_line il

inner join invoice i on il.invoice\_id = i.invoice\_id

inner join customer c on i.customer\_id = c.customer\_id

inner join track t on il.track\_id = t.track\_id

inner join album al on t.album\_id = al.album\_id

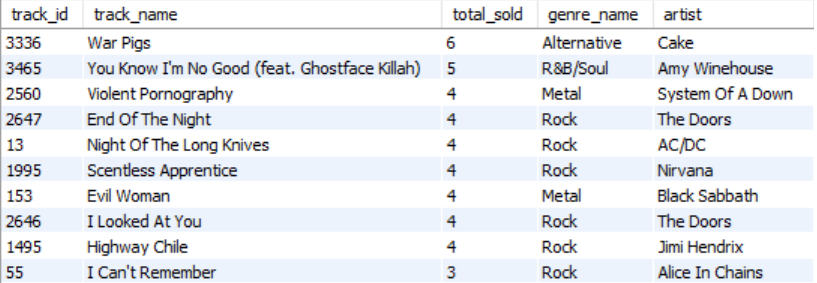
inner join artist a on al.artist\_id = a.artist\_id

inner join genre g on t.genre\_id = g.genre\_id

where c.country = 'USA'

group by t.track\_id, t.name, g.name, a.name

order by total\_sold desc limit 10;



Top-selling tracks in USA is **War Pigs** with the genre **Alternative** and the artist is **Cake**

The Most Famous Genre is **Rock** where the top-most artist is **The Doors** with the track name as **End of the Night** and **I Looked at You**

--Top Artist in the USA

select a.artist\_id, a.name as artist\_name, sum(il.quantity) as total\_quantities\_sold from invoice\_line il

inner join invoice i on il.invoice\_id = i.invoice\_id

inner join customer c on i.customer\_id = c.customer\_id

inner join track t on il.track\_id = t.track\_id

inner join album al on t.album\_id = al.album\_id

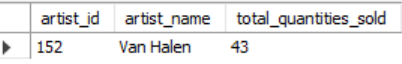
inner join artist a on al.artist\_id = a.artist\_id

inner join genre g on t.genre\_id = g.genre\_id

where c.country = 'USA'

group by a.artist\_id, a.name

order by total\_quantities\_sold desc limit 1;



The Van Halen is one of the top most leading artist in USA with around 43 quantities sold

--Most Famous Genre in USA

select g.genre\_id, ar.artist\_id,g.name as genre\_name, sum(il.quantity) as total\_quantities\_sold from

invoice\_line il

inner join invoice i on il.invoice\_id = i.invoice\_id

inner join customer\_view c on c.customer\_id = i.customer\_id

inner join track t on il.track\_id = t.track\_id

inner join album a on t.album\_id = a.album\_id

inner join artist ar on ar.artist\_id =a.artist\_id

inner join genre g on g.genre\_id= t.genre\_id

where c.country = 'USA' and ar.artist\_id = 152

group by g.name, g.genre\_id, ar.artist\_id

order by total\_quantities\_sold desc;



As mentioned previously the **Rock** genre is most famous genre in USA

1. What is the customer demographic breakdown (age, gender, location) of Chinook's customer base?

select country, coalesce(state,"NA") as state, city, count(\*) as total\_customers from customer

group by country, coalesce(state,"NA"), city

order by total\_customers desc,country;



1. Calculate the total revenue and number of invoices for each country, state, and city:

I have used Customer\_view which was created initially to remove the null values

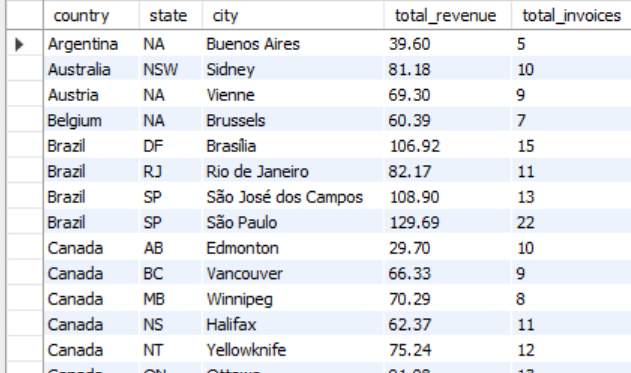
select c.country, c.state, c.city, sum(i.total) as total\_revenue, count(invoice\_id) as total\_invoices

from customer\_view c inner join invoice i

on i.customer\_id = c.customer\_id

group by c.country,c.state, c.city

order by country;



1. Find the top 5 customers by total revenue in each country

with cte as (select c.customer\_id, c.first\_name, c.last\_name, c.country, sum(i.total) as total\_revenue,

dense\_rank() over(partition by c.country order by sum(i.total) desc) as rnk

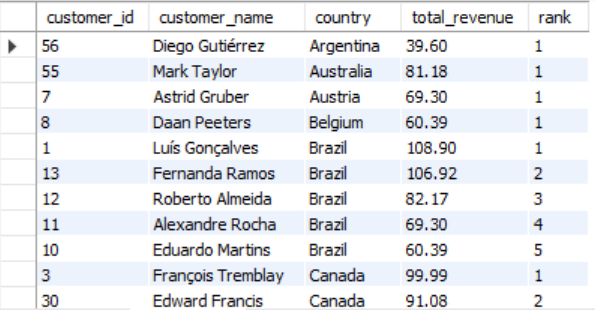
from customer\_view c inner join invoice i

on i.customer\_id = c.customer\_id

group by c.customer\_id, c.country

order by c.country, rnk)

select customer\_id, concat(first\_name," ", last\_name) as customer\_name, country, total\_revenue, rnk as "rank" from cte where rnk between 1 and 5;



1. Identify the top-selling track for each customer

select c.customer\_id, concat(first\_name, " ",last\_name) as full\_name, t.name as track\_name, sum(quantity) as total\_quantities from customer c

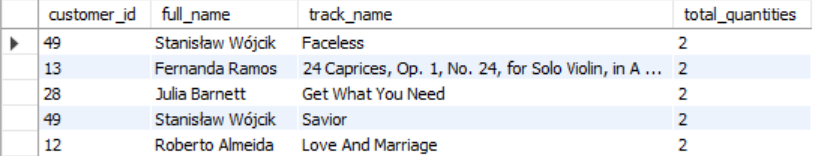
inner join invoice i on i.customer\_id = c.customer\_id

inner join invoice\_line il on il.invoice\_id = i.invoice\_id

inner join track t on t.track\_id = il.track\_id

group by c.customer\_id, concat(first\_name, " ",last\_name), t.name

order by total\_quantities desc;



Frantisek Wichterlova is the top most customer from city Prague has purchased highest no. of quantities 146 overall considering all the tracks

1. Are there any patterns or trends in customer purchasing behavior (e.g., frequency of purchases, preferred payment methods, average order value)?

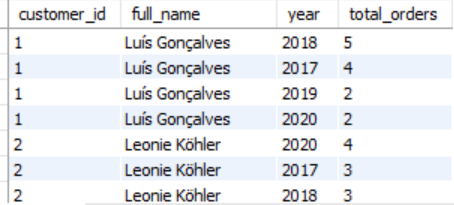
-Analysing the Frequency orders placed by each customer Year-on-Year

select c.customer\_id,concat(first\_name, " ",last\_name) as full\_name, year(invoice\_date)as "year",

count(invoice\_id) as total\_orders

from customer\_view c inner join invoice i

on i.customer\_id=c.customer\_id

group by c.customer\_id, concat(first\_name, " ",last\_name),year(invoice\_date) order by total\_orders desc;

* Average total value spent by each customer in each year is

select c.customer\_id,concat(first\_name, " ",last\_name) as full\_name, year(invoice\_date)as "year",

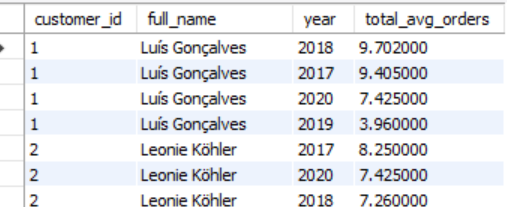
avg(total) as total\_avg\_orders

from customer\_view c inner join invoice i

on i.customer\_id=c.customer\_id

group by c.customer\_id, concat(first\_name, " ",last\_name),year(invoice\_date)

order by customer\_id, total\_avg\_orders desc;



1. What is the customer churn rate?

with cte as (select c.customer\_id, invoice\_date from customer\_view c inner join invoice I on i.customer\_id=c.customer\_id),

churn as (select year(invoice\_date) as "year", count(distinct customer\_id)as no\_of\_customer,

lag(count(distinct customer\_id)) over(order by year(invoice\_date)) as prev\_yr\_count

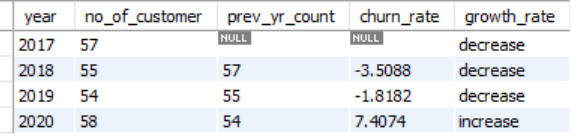
from cte group by year(invoice\_date))

select year, no\_of\_customer, prev\_yr\_count, ((no\_of\_customer-prev\_yr\_count)\*100/prev\_yr\_count) as churn\_rate,

case when no\_of\_customer> prev\_yr\_count then "increase"

else "decrease" end as growth\_rate

from churn;



-Though churn\_rate is very less and minimal and the highest churn\_rate between 2017 and 2020 is 2018 with highest churn\_rate as -3.5 where we have seen some decrease in no\_of\_customer during these years

In 2020, there is increase in customers compared to the previous year.

1. Calculate the percentage of total sales contributed by each genre in the USA and identify the best-selling genres and artists.

- Percentage of total sales contributed by each genre in the USA

with cte\_usa as (select t.genre\_id, g.name as genre\_name,

sum(i.quantity \* i.unit\_price) as total\_genre\_price, ic.billing\_country as country

from invoice\_line i inner join

track t on t.track\_id = i.track\_id

inner join invoice ic on ic.invoice\_id = i.invoice\_id

inner join genre g on g.genre\_id = t.genre\_id

where ic.billing\_country = "USA"

group by t.genre\_id, ic.billing\_country,g.name

order by total\_genre\_price desc),

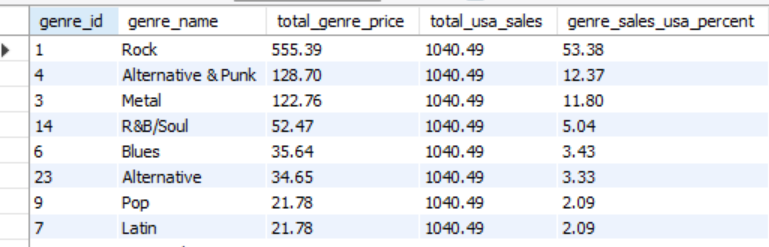
cte\_usa1 as ( select country, sum(total\_genre\_price) as total\_usa\_sales from cte\_usa group by country)

select genre\_id, genre\_name, total\_genre\_price, total\_usa\_sales,

round((total\_genre\_price/ total\_usa\_sales)\*100,2) as genre\_sales\_usa\_percent

from cte\_usa cross join cte\_usa1

order by genre\_sales\_usa\_percent desc;



Identifying the best-selling genres and artists

with cte\_usa as (select t.genre\_id, g.name as genre\_name,

sum(i.quantity \* i.unit\_price) as total\_genre\_price, ic.billing\_country as country

from invoice\_line i inner join

track t on t.track\_id = i.track\_id

inner join invoice ic on ic.invoice\_id = i.invoice\_id

inner join genre g on g.genre\_id = t.genre\_id

where ic.billing\_country = "USA"

group by t.genre\_id, ic.billing\_country,g.name

order by total\_genre\_price desc),

cte\_usa1 as ( select country, sum(total\_genre\_price) as total\_usa\_sales from cte\_usa group by country)

select genre\_id, genre\_name, total\_genre\_price, total\_usa\_sales,

round((total\_genre\_price/ total\_usa\_sales)\*100,2) as genre\_sales\_usa\_percent

from cte\_usa cross join cte\_usa1

order by genre\_sales\_usa\_percent desc;

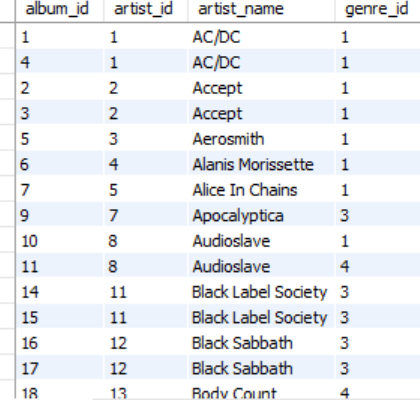


select distinct t.album\_id, a.artist\_id, ar.name as artist\_name, t.genre\_id from track t

inner join album a on a.album\_id = t.album\_id

inner join artist ar on ar.artist\_id = a.artist\_id where t.genre\_id in (1,4,3)

order by artist\_i;



-Based on the genre the total price of genre in USA are from Rock, Alternative & Punk, Metal which are genre\_id is 1,3,4 with percentage of

53.38% in Rock genre, 12.37% in Alternative & Punk and 11.80 in Metal genres respectively.

1. Find customers who have purchased tracks from at least 3 different genres

with cte as ( select c.customer\_id, concat(first\_name, " ",last\_name ) as full\_name, t.track\_id, g.genre\_id from customer c

inner join invoice i on c.customer\_id = i.customer\_id

inner join invoice\_line il on il.invoice\_id = i.invoice\_id

inner join track t on t.track\_id = il.track\_id

inner join genre g on g.genre\_id = t.genre\_id )

select customer\_id, full\_name, count(distinct genre\_id) as no\_of\_genre from cte group by customer\_id, full\_name having count(distinct genre\_id) >=3;



-As required got the desired output of customers where they have purchased more than 3 genres in overall.

1. Rank genres based on their sales performance in the USA

with cte as (select g.genre\_id, g.name as genre\_name,

sum(il.unit\_price \* il.quantity) as total\_price, i.billing\_country as country from invoice\_line il

inner join track t on t.track\_id = il.track\_id

inner join genre g on g.genre\_id=t.genre\_id

inner join invoice i on i.invoice\_id = il.invoice\_id

where i.billing\_country= "USA"

group by t.track\_id, g.genre\_id, i.billing\_country),

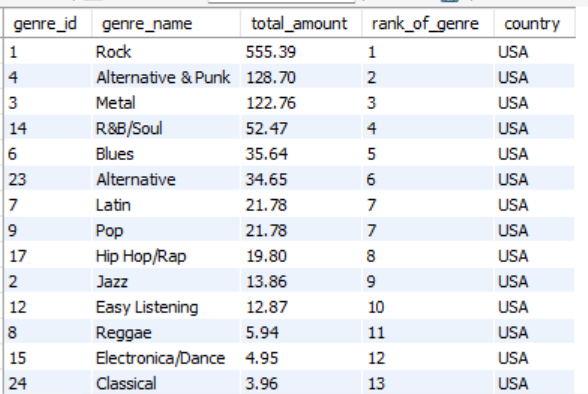
cte1 as (select genre\_id, genre\_name, country, sum(total\_price) as total\_amount,

dense\_rank() over(order by sum(total\_price) desc) as rank\_of\_genre

from cte

group by genre\_id)

select genre\_id, genre\_name, total\_amount, rank\_of\_genre, country from cte1;



In USA the ranks are distributed among the genres, I have used the dense\_rank option as I do not want to skip the ranks between the genres

The highest rank\_of genre for genre in USA is Rock as it have total amount of 555.39 with highest sale

1. Identify customers who have not made a purchase in the last 3 months

with last\_date as

( select max(invoice\_date) as max\_invoice\_date, customer\_id from invoice group by customer\_id),

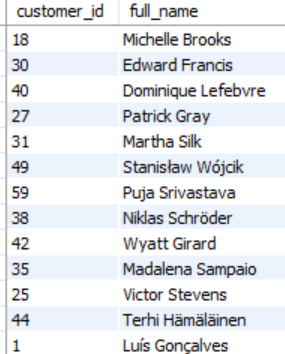
customers\_purchases as (select distinct c.customer\_id, concat(first\_name, " ", last\_name) as full\_name from customer c

inner join last\_date l on

l.customer\_id = c.customer\_id inner join invoice i on i.customer\_id = c.customer\_id

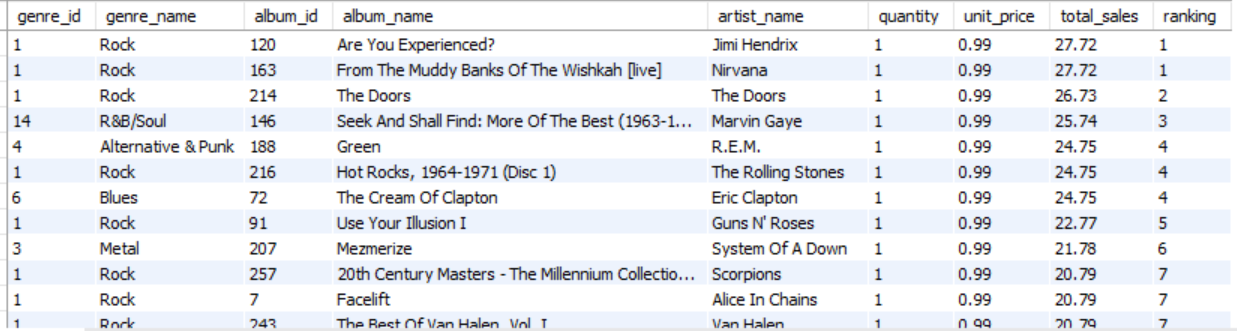
where i.invoice\_date < date\_sub(l.max\_invoice\_date, interval 3 month)or i.invoice\_date is null)

select customer\_id, full\_name from customers\_purchases;



Subjective Questions

1. Recommend the three albums from the new record label that should be prioritised for advertising and promotion in the USA based on genre sales analysis.



The three most popular and albums in USA based on the quantities sold was

1. Are you Experienced? - made by Pink Floyd & From the Muddy Banks of the Wishkah [live] – made by Corinne Bailey Rae were sold highest number of quantities which are 28 in USA in the Rock genre

2. The Doors – made by Academy of St. Martin in the Fields & Sir Neville Marriner was standing in second place with 27 quantities were sold across USA

3. Seek And Shall Find: More Of The Best (1963-1981) made by Titas

was standing in second place with 26 quantities were sold across USA

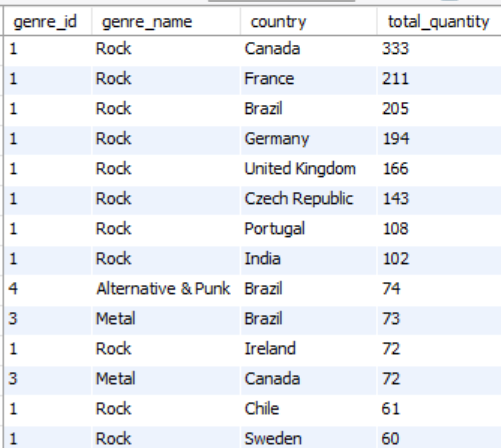
**Insights:**

* **Genre Popularity**: Rock is the dominant genre with the highest total sales figures across all the albums. Most of the albums belongs to this genre, indicating that this Rock genre has a strong market preference for music in the USA.
* **Sales Figures**: The selected albums have the highest total genre sales with highest number of copies sold across the USA, making them strong candidates for promotions.

1. Determine the top-selling genres in countries other than the USA and identify any commonalities or differences.

Rock genre is most popular genre across the regions irrespective of USA or other regions

* These are the top 5 genres across the regions where
* 1. Rock
* 2. Metal
* 3. Alternative & Punk
* 4. Latin
* 5. Jazz



**commonalities:**

As based on analysis the Rock genre is top most favourite genre for the most of the countries except some countries which are vibing to the other genres.

**differences:**

Based on above analysis the top most genre in Brazil is **Alternative & Punk** and second followed by **Metal** in Brazil. It is showing a more varied musical taste compared to other countries.

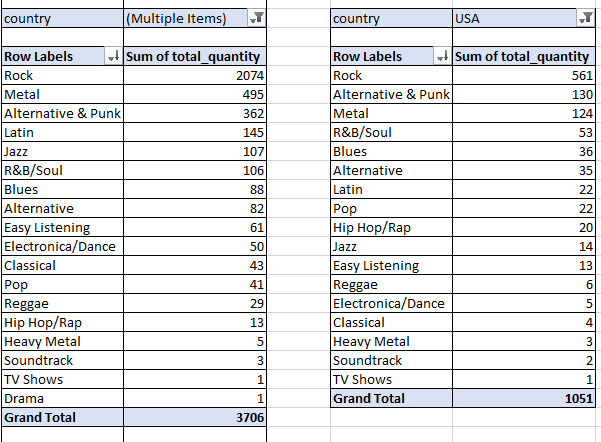
This shows that if any promotion or concert is happening in Brazil, they can perform the either **Alternative & Punk** or **Metal.**

Apart from Rock genre, the Metal genre has pretty good record in Canada & Brazil which shows that those customers have a varied music taste and shown interest in different genres as they are showing the versatility in the music taste.

**Insights:**

* Rock is the overwhelmingly dominant genre across the most of the countries, indicating a universal appeal for the genre as many are celebrating this.
* Brazil has a more diverse taste with having the strong sales in Alternative & Punk and Metal followed by the Rock.
* The preference for Metal is notably strong in Canada and Brazil, which differs from other countries where Rock dominates in most of the countries.

Based on the analysis shows that while Rock is a global favourite for most of the parts across the globe whereas the countries like Brazil which exhibits the varied musical taste which could be important for the artists to keep engaging the audience through marketing, concerts and promotion.



1. Customer Purchasing Behaviour Analysis: How do the purchasing habits (frequency, basket size, spending amount) of long-term customers differ from those of new customers? What insights can these patterns provide about customer loyalty and retention strategies?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Row Labels** | **Average of purchase\_frequency** | **Average of avg\_basket\_size** | **Average of avg\_spending\_amount** | **Average of date\_diff** |
| long term | 10.6 | 1.0 | 9.8 | 1259.9 |
| new | 6.0 | 1.0 | 9.6 | 799.5 |
| **Grand Total** | **10.4** | **1.0** | **9.8** | **1244.3** |

with customerinvoicedates as (

select

c.customer\_id,c.first\_name, c.last\_name,

min(date(i.invoice\_date)) as first\_purchase\_date,

max(date(i.invoice\_date)) as last\_purchase\_date,

count(distinct i.invoice\_id) as purchase\_frequency,

round(avg(il.quantity),2) as avg\_basket\_size,

round(avg(i.total),2) as avg\_spending\_amount

from customer c

join invoice i on c.customer\_id = i.customer\_id

join invoice\_line il on i.invoice\_id = il.invoice\_id

group by c.customer\_id,c.first\_name, c.last\_name),

customercategory as (

select

\*,

datediff(last\_purchase\_date,first\_purchase\_date) as date\_diff,

case when datediff(last\_purchase\_date,first\_purchase\_date) > 910 then 'long term' else 'new' end as category\_type

from customerinvoicedates)

select \* from customercategory order by customer\_id;

**Insights:**

Assuming the customer hasn’t made any purchases over 2.5 years, categorizing them as New as these users are not a regular user and the customers who made purchases for every 2.5 years are being categorized as the Long-Term Users

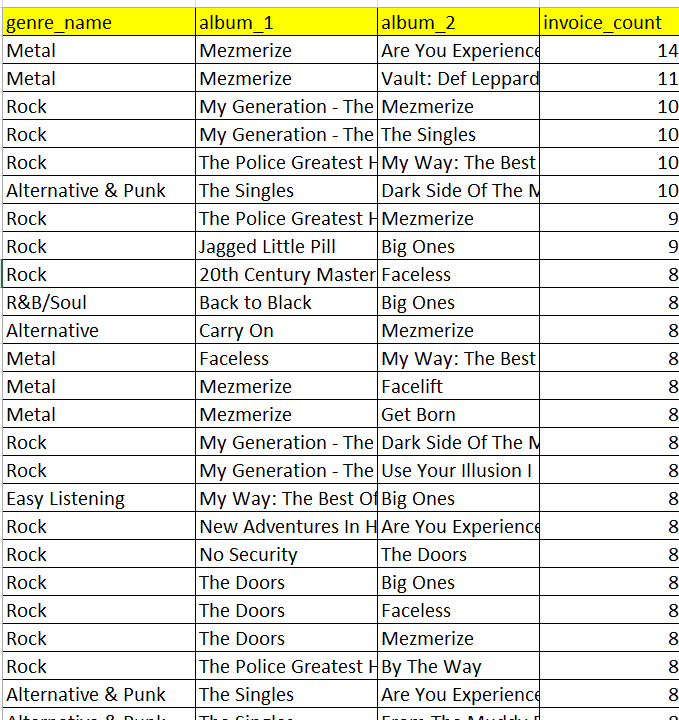
As we can see many customers are falling under Long-Term category which shows the loyalty and shows that retention rate is high

We can see a huge gap between the purchase frequency which shows how many purchases are being made by both categories and the majority of customers falling New category made the purchases below 7 on an overall average.

**Recommendation:**

* We can start rolling out some offers to the customers falling under New category or annual packages with offers to convert them into Long-Term Category
* For people who are already in Long-Term can be given good offers to purchase and provide some support to be retain as Long-Term Category
* Referrals can be used and once the reference is used some membership points can be provided for both old & new customers to keep them engaging by adding some perks & benefits which can be used for next purchase or any coupons for shopping can be used.

1. Product Affinity Analysis: Which music genres, artists, or albums are frequently purchased together by customers? How can this information guide product recommendations and cross-selling initiatives?



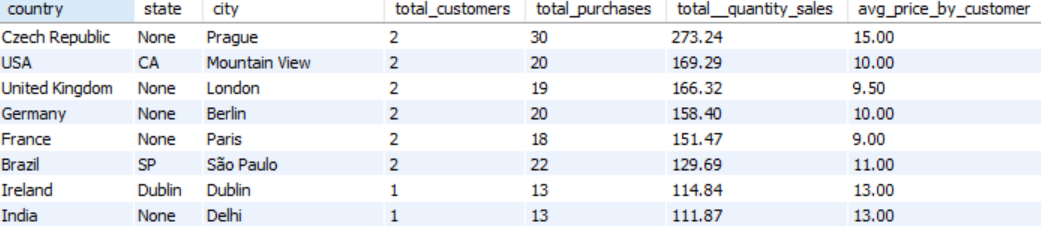
**Insights:**

* The Rock genre has been purchased by most of the customers along with the other genres with the purchase counts
* Where Mezmerize is top most album purchased with popular album paired with several others combination of another album

**Recommendation:**

* We can start promoting the Rock & Metal in the ads, banners or start promoting these genres while purchasing the albums.
* Special Editions: Create box sets with frequently paired albums to attract collectors with some unique offers.
* Curated Collections: Create and market "Perfect Pairings" or "Ultimate Combo Packs" based on genre and artist affinities.
* Personalized Recommendations: Suggest related albums (e.g., Metal or Alternative & Punk) when a customer buys a Rock album.

1. Regional Market Analysis: Do customer purchasing behaviours and churn rates vary across different geographic regions or store locations? How might these correlate with local demographic or economic factors?



Customer Analysis:

* From Czech Republic the total purchases made are 30 which is highest across all the regions and revenue generated with highest of $273.24 with the average orders by each customer is 15 in number
* As lower average through the regions is lying in Edmonton in Canada with $29.70 were generated followed by Copenhagen from Denmark generated second lowest revenue which is $29.70
* As sales are concerned the customers are not purchasing the items which is correlated to the revenue

**Insights:**

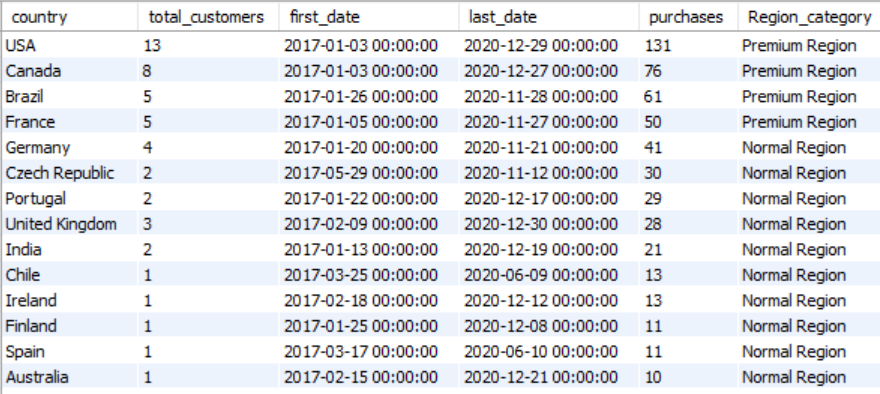
* The Top 5 Cities which most revenue can be treated as Premium where they are generating the highest sale revenue can be categorized as **Affluent** regions
* The Cities with Moderate revenue generated by customers can be categorized as **Moderate** regions
* The Cities with Low revenue generated by customers can be categorized as **Low** regions

**Churn Rates**

* The Regions which are having greater than equal to 5 customers has low churn rate as it is having more than 5 customers as the revenue generation won’t be stopped and it needs to be taken care as the region falls under the Premium Category
* The Regions which are having less than or equal to 3 customers will have the less churn rate as if customer is not purchasing any items, then the revenue won’t be generated from that region no more and is correlated to each other.

**Recommendation**

* For the premium regions there should be separate discounts should be provided, if needed annual packs with discount can be provided as they are generating more revenue.
* For the regions with only 2 or less customers should be retained by providing multiple offers and referral bonus also can be provided as it will help to gain the customers.
* The content should be delivered based on the customers interest across the regions where we are having the less customers by asking the survey and by following the current trends the promotions can be done to engage the customers.

****

1. Customer Risk Profiling: Based on customer profiles (age, gender, location, purchase history), which customer segments are more likely to churn or pose a higher risk of reduced spending? What factors contribute to this risk?

select i.billing\_country as country, i.billing\_state as state, i.billing\_city as city, count(distinct c.customer\_id) as total\_customers,

count( distinct i.invoice\_id) as purchase\_orders, round(count(distinct i.invoice\_id)/count(distinct c.customer\_id),2) as avg\_cust\_orders,

round(avg(i.total),2) as avg\_spending\_amount,

case when round(count(distinct i.invoice\_id)/count(distinct c.customer\_id),2) >10 then "Low Risk"

else "High Risk" end as Risk\_category

from customer c left join invoice i

on c.customer\_id = i.customer\_id

inner join invoice\_line il on il.invoice\_id = i.invoice\_id

group by i.billing\_country , i.billing\_state , i.billing\_city

order by purchase\_orders desc;



We have categorized the Risk Category based on the average orders placed by each customer in the country with respect to their state and respective city.

**Insights**

* **High Spending with High Frequency (Potential Churn Risk):**
* There is a high potential risk of these below regions as though there is highest revenue generated but if customer stops renewing or purchasing the store will see a gradual fall as these regions generate high revenue.
* Regions: Prague, Czech Republic; Delhi, India; Brasília, Brazil.
* Risk Factors: The High expectations and increased competition may lead to churn if service or deals decline.
* **Moderate Spending with Moderate Frequency:**
* Regions: Mountain View, USA; Stuttgart, Germany.
* Risk Factors: Dependence on satisfaction and economic sensitivity could lead to reduced spending or churn if conditions change.
* **Low Spending with High Frequency:**
* Regions: Yellowknife, Canada; Stockholm, Sweden.
* Risk Factors: Value-seeking and promotional dependence make these customers sensitive to price increases or reduced deals.
* **Low Spending with Low Frequency (High Churn Risk):**
* Regions: Edmonton, Canada; Copenhagen, Denmark.
* Risk Factors: Low engagement and price sensitivity increase churn risk if offerings aren't compelling or prices rise.

**Recommendations**

**Measures on High Risk:**

Product Quality, Content Moderation, Competitive marketplaces, Customer Service.

**Enhanced Engagement:** Use loyalty programs for every referral, partnership with other agencies for vouches and personalized annual, quarterly or monthly offers for high-spending customers.

**Marketing:**

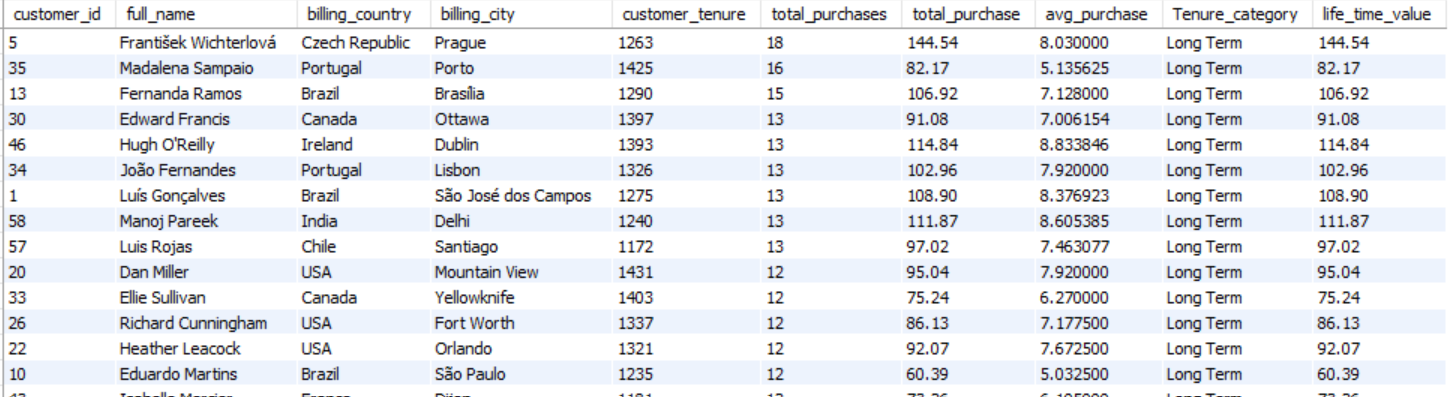
Promote the content based on the regions and highlight the value for money by maintaining regular events, concerts, promotions for high revenue areas

**Churn Predictions:**

Implement predictive analytics by automation to identify and address the risk-based customers and regions early based on the customers usage.

These strategies will help reduce churn and maximize customer lifetime value.

1. Customer Lifetime Value Modelling: How can you leverage customer data (tenure, purchase history, engagement) to predict the lifetime value of different customer segments? This could inform targeted marketing and loyalty program strategies. Can you observe any common characteristics or purchase patterns among customers who have stopped purchasing?



we can analyse the customer lifetime value (CLV) with above part.

With respect to country and categorized the customers as Long-Term or Short-Term based on the tenure which means the difference the first purchase date and last purchase date if customer is staying more than 1000 days and making the purchases marked them as Long-Term else as Short-Term

|  |  |  |
| --- | --- | --- |
| **Row Labels** | **Count of customer\_id** | **Sum of life\_time\_value** |
| Long Term | 55 | 4460.94 |
| Short Term | 4 | 248.49 |
| **Grand Total** | **59** | **4709.43** |

* As Short Term only 4 customers are concerned Root Cause Analysis can be performed to know the reason for this Churn to improve the Customers by decreasing the Churn rate
* As there is not much revenue generated from Short-Term but compared to Long-Term the average per customer is major difference which was made by customer during the purchases
* Once RCA is done then the content can be delivered as needed or requested by taking necessary actions to avoid in those regions by not spending too much on the Short-Term region customers as we don’t have a guarantee of customers retention.
* Customers with lower average order values and having long tenure is not a good sign as they can be categorized as occasional customer which may be at risk if they don’t purchase any order value as though they are Long-Term and not purchasing by having no part in sales

1. If data on promotional campaigns (discounts, events, email marketing) is available, how could you measure their impact on customer acquisition, retention, and overall sales?

As of now we don’t have the data of promotional campaigns we cannot give accurate information

If the data exists then we can analyse the data by these aspects

**Promotional Campaigns:**

* Begin by analyzing the dataset to determine the specific dates or timeframes when promotional campaigns were active.

**Data Segmentation**

* Divide the dataset into two segments: one representing periods with active promotions, and the other capturing periods without any promotions.
* During the active promotions divide the dataset into 2 sub segments as one is during the promotion campaign and next is before 3 days & after 3 days of the promotion campaign to check the buzz that is being created by the promotion campaigns

**Customer Acquisition Analysis**

* Evaluate the number of new customers acquired during promotional periods and compare it to the acquisition rate during non-promotional periods to monitor the promotion campaign whether it is success or not.

**Customer Retention Assessment**

* Examine customer retention by tracking repeat purchases made during and after the promotional campaigns to assess long-term engagement.

**Sales Performance Comparison**

* Analyse the key sales metrics such as total revenue, average order value, and basket size across both promotional and non-promotional periods to measure campaign effectiveness.

1. How would you approach this problem, if the objective and subjective questions weren't given?

* Based on the dataset we can follow these steps

1. Extract & Load the data to check the which type of data is existing
2. Clean the data by filling the NULL Values using Coalesce functions or ifnull functions
3. Create the View for the cleaned data tables to reuse the data tables for the future purposes.
4. Check the Primary Key & Foreign Key for the references as we have to check how the tables relate to each other with the relations
5. Start working on Data by checking the Customer information which includes age, country etc.., the favourite genres, top most albums, top artists to get the top most genres across the regions.
6. Identify the patterns of sales in the data with respect to genres, artists, albums by the customers
7. Check the longevity of the customers whether they are Short-Term or Long-Term
8. Make sure to find the purchase patterns by the customer during the Promotional or Non-Promotional Campaigns
9. Use predictive analysis for the promotions by checking the most buzzing genres across the regions
10. Identify the average order value by each customer with respect to average order value across the region with the number of customers in each category.
11. Provide the Strategic Insights & Recommendation
12. Summarize the above needy information and present it in a report or dashboard for the better understanding

This approach focuses on in-depth data analysis to generate actionable insights that support strategic decision-making, improve customer retention, boost sales performance, and uncover areas for growth. It is a continuous, evolving process that adapts as new patterns and findings emerge

1. How can you alter the "Albums" table to add a new column named "ReleaseYear" of type INTEGER to store the release year of each album?

Use alter to add a new column to the table and then update the values as required in the new column or attribute

alter table album

add column ReleaseYear int;

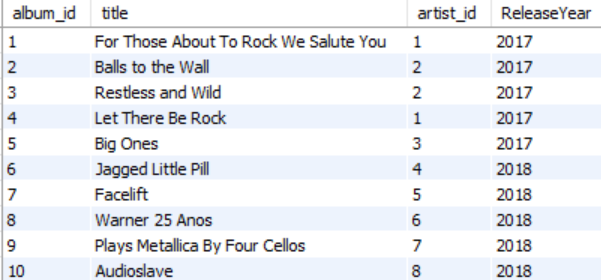
update album set ReleaseYear = 2017

where album\_id in (1,2,3,4,5);

update album set ReleaseYear = 2018

where album\_id in (6,7,8,9,10);

select \* from album;



1. Chinook is interested in understanding the purchasing behavior of customers based on their geographical location. They want to know the average total amount spent by customers from each country, along with the number of customers and the average number of tracks purchased per customer. Write an SQL query to provide this information.

with tracks\_by\_customer as (

select i.customer\_id, sum(il.quantity) as total\_tracks

from invoice i

inner join invoice\_line il on i.invoice\_id = il.invoice\_id

group by i.customer\_id

),

customer\_spent as (

select c.country, c.customer\_id, sum(i.total) as total\_spent, tpc.total\_tracks

from customer c

inner join invoice i on c.customer\_id = i.customer\_id

inner join tracks\_by\_customer tpc on c.customer\_id = tpc.customer\_id

group by c.country, c.customer\_id, tpc.total\_tracks

)

select cs.country,

count(distinct cs.customer\_id) as number\_of\_customers,

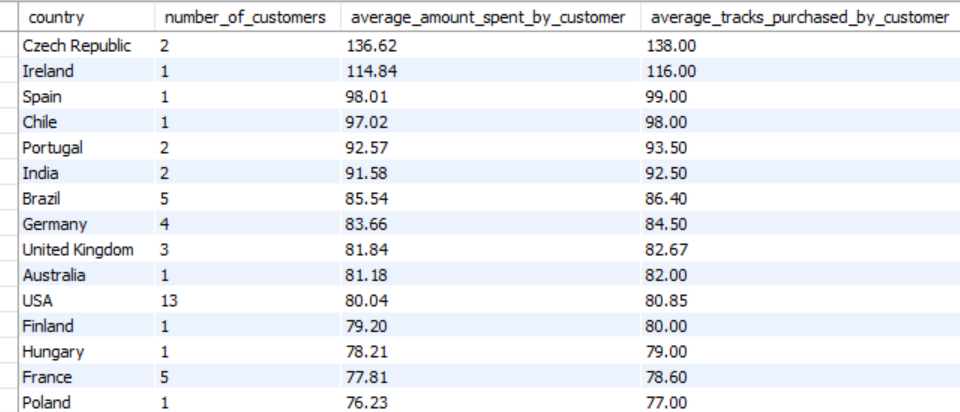
round(avg(cs.total\_spent),2) as average\_amount\_spent\_by\_customer,

round(avg(cs.total\_tracks),2) as average\_tracks\_purchased\_by\_customer

from customer\_spent cs

group by cs.country

order by average\_amount\_spent\_by\_customer desc;



THE END