Objective Questions

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

**Ans-:** **The provided SQL database is well-structured and free of duplicate values, as confirmed by running appropriate queries. However, there are missing values in three tables:**

1. **Customer Table:**
   * **47 fax values are null.**
   * **29 state values are null.**
   * **49 company values are null.**
2. **Employee Table:**
   * **The reports\_to field is null for employee\_id = 1.**
3. **Track Table:**
   * **978 composer values are null.**

**These null values can be effectively handled using the COALESCE function in SQL to provide default values or alternative representations where needed.**

**My SQl Query-: -- first in album table**

**SELECT title, artist\_id, COUNT(\*) AS duplicate\_count**

**FROM album**

**GROUP BY title, artist\_id**

**HAVING COUNT(\*) > 1;**

**-- in artist table**

**SELECT \* FROM artist WHERE name IS NULL; -- for missing values**

**SELECT name, COUNT(\*) FROM artist GROUP BY name HAVING COUNT(\*) > 1; -- for duplicate value**

**-- in customer table**

**SELECT \* FROM customer WHERE address IS NULL OR city IS NULL; -- for missing values**

**SELECT email, COUNT(\*) FROM customer GROUP BY email HAVING COUNT(\*) > 1; -- for duplicate value**

**-- in employee Table**

**SELECT \* FROM employee WHERE reports\_to IS NULL; -- for missing values**

**UPDATE employee SET reports\_to = 0 WHERE reports\_to IS NULL; -- update missing value**

**SELECT email, COUNT(\*) FROM employee GROUP BY email HAVING COUNT(\*) > 1; -- for duplicate value**

**-- in genre table**

**SELECT \* FROM genre WHERE name IS NULL;**

**UPDATE genre SET name = 'Unknown' WHERE name IS NULL; -- for missing value**

**SELECT name, COUNT(\*) FROM genre GROUP BY name HAVING COUNT(\*) > 1; -- for duplicate value**

**-- in invoice table**

**SELECT \* FROM invoice WHERE billing\_address IS NULL; -- for missing value**

**-- invoice\_line Table**

**SELECT invoice\_id, track\_id, COUNT(\*)**

**FROM invoice\_line**

**GROUP BY invoice\_id, track\_id**

**HAVING COUNT(\*) > 1; -- for duplicate value**

**select \* from invoice\_line;**

**-- there are having the duplicate number because from one invoice\_id and from one track\_id**

**-- there we can track multiple purchese from one track\_id and there can be multiple purchese from one invoice\_id**

**-- media\_type Table**

**SELECT \* FROM media\_type WHERE name IS NULL; -- for missing value**

**SELECT name, COUNT(\*) FROM media\_type GROUP BY name HAVING COUNT(\*) > 1; -- for duplicate value**

**-- playlist Table**

**SELECT \* FROM playlist WHERE name IS NULL; -- for missing value**

**SELECT name, COUNT(\*) FROM playlist GROUP BY name HAVING COUNT(\*) > 1; -- for duplicate value**

**-- playlist\_track Table**

**SELECT playlist\_id, track\_id, COUNT(\*)**

**FROM playlist\_track**

**GROUP BY playlist\_id, track\_id**

**HAVING COUNT(\*) > 1; -- for duplicate value**

**-- track Table**

**SELECT \* FROM track WHERE album\_id IS NULL;**

**SELECT name, album\_id, media\_type\_id, COUNT(\*)**

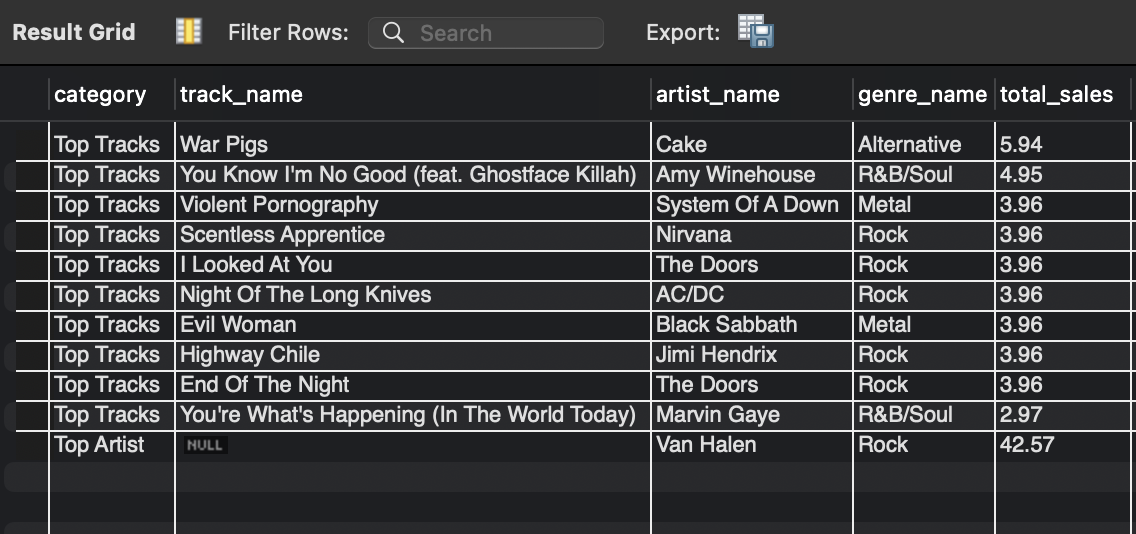
**FROM track**

**GROUP BY name, album\_id, media\_type\_id**

**HAVING COUNT(\*) > 1;**

2 . Find the top-selling tracks and top artists in the USA and identify their most famous genres.

**Ans-: Below are the top 10 selling Tracks of the USA, War Pigs by Cake being on top**.



**Below are the top most preferred genre in USA by number of records sold:**

**Rock is the most preferred genre in the USA.**

**My SQL query-: SELECT i.invoice\_id**

**FROM invoice i**

**WHERE i.billing\_country = 'USA';**

**-- Step 1: Find invoices in the USA**

**WITH usa\_invoices AS (**

**SELECT i.invoice\_id**

**FROM invoice i**

**WHERE i.billing\_country = 'USA'**

**),**

**-- Step 2: Calculate total sales per track for USA invoices**

**track\_sales AS (**

**SELECT**

**il.track\_id,**

**SUM(il.unit\_price \* il.quantity) AS total\_sales**

**FROM invoice\_line il**

**JOIN usa\_invoices ui ON il.invoice\_id = ui.invoice\_id**

**GROUP BY il.track\_id**

**),**

**-- Step 3: Get top-selling tracks and their details**

**top\_tracks AS (**

**SELECT**

**t.track\_id,**

**t.name AS track\_name,**

**a.artist\_id,**

**ar.name AS artist\_name,**

**g.genre\_id,**

**g.name AS genre\_name,**

**ts.total\_sales**

**FROM track\_sales ts**

**JOIN track t ON ts.track\_id = t.track\_id**

**JOIN album a ON t.album\_id = a.album\_id**

**JOIN artist ar ON a.artist\_id = ar.artist\_id**

**JOIN genre g ON t.genre\_id = g.genre\_id**

**ORDER BY ts.total\_sales DESC**

**LIMIT 10 -- top 10 tracks**

**),**

**-- Step 4: Identify the top artist and their most popular genre in the USA**

**top\_artist\_genre AS (**

**SELECT**

**ar.artist\_id,**

**ar.name AS artist\_name,**

**g.genre\_id,**

**g.name AS genre\_name,**

**SUM(ts.total\_sales) AS artist\_total\_sales**

**FROM track\_sales ts**

**JOIN track t ON ts.track\_id = t.track\_id**

**JOIN album a ON t.album\_id = a.album\_id**

**JOIN artist ar ON a.artist\_id = ar.artist\_id**

**JOIN genre g ON t.genre\_id = g.genre\_id**

**GROUP BY ar.artist\_id, g.genre\_id**

**ORDER BY artist\_total\_sales DESC**

**LIMIT 1 -- top artist**

**)**

**-- Step 5: Display the results**

**SELECT**

**'Top Tracks' AS category,**

**tt.track\_name,**

**tt.artist\_name,**

**tt.genre\_name,**

**tt.total\_sales**

**FROM top\_tracks tt**

**UNION ALL**

**SELECT**

**'Top Artist' AS category,**

**NULL AS track\_name,**

**tg.artist\_name,**

**tg.genre\_name,**

**tg.artist\_total\_sales AS total\_sales**

**FROM top\_artist\_genre tg;**

**-- finding top selling genre in USA**

**select Top\_Genre from**

**(**

**select g.name as Top\_Genre**

**from track t**

**left join invoice\_line il on il.track\_id = t.track\_id**

**left join invoice i on i.invoice\_id = il.invoice\_id**

**left join genre g on t.genre\_id = g.genre\_id**

**where i.billing\_country = 'USA'**

**group by g.name**

**order by sum(il.quantity) desc**

**limit 10**

**) sub\_table;**

**Insights**

* **Rock dominates US sales.**
* **"War Pigs" by Cake is the top track.**

**Recommendations**

* **Prioritize Rock in inventory and marketing.**
* **Use genre-specific promotions.**
* **Consider long-term genre diversification.**

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

**Ans-: The Chinook music store serves customers in 24 countries, with the USA leading as the country with the highest number of customers, totaling 13.**

**In data we don't have any table or column where we have the age and gender. So we can’t analyze by age and gender.**

**My SQL Query-: select \* from customer;**

**SELECT**

**country,**

**COUNT(customer\_id) AS customer\_count**

**FROM customer**

**GROUP BY country**

**ORDER BY customer\_count DESC;**

**Insights:**

* **Customers are distributed across 24 countries.**
* **The USA has the highest number of customers.**

**Recommendations:**

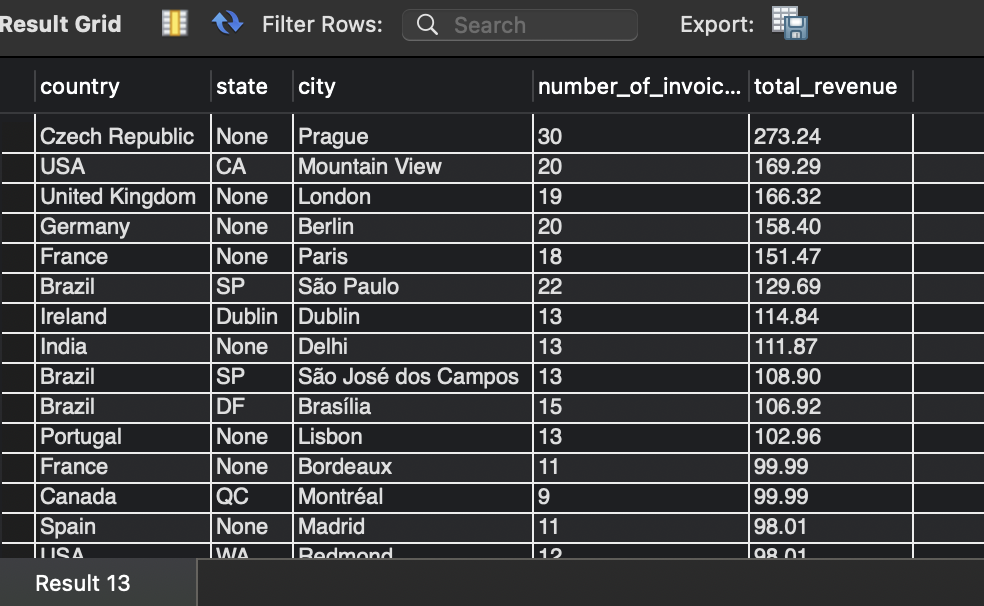
1. **Prioritize the US market, leveraging its significant customer base.**
2. **Identify growth potential in key countries like Canada, Brazil, France, and Germany.**
3. **Develop region-specific marketing strategies to enhance engagement and expand the customer base.**

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

**Ans-: Insights:**

* **Prague (Czech Republic) leads in revenue generation.**
* **São Paulo (Brazil) has a high volume of invoices but comparatively lower revenue.**

**Recommendations:**

1. **Explore the factors contributing to Prague's high revenue per invoice and apply these insights to other cities.**
2. **Conduct a detailed analysis of revenue per invoice across all locations to pinpoint opportunities for improvement and refine pricing strategies.**

My Sql Query-: SELECT

billing\_country AS country,

billing\_state AS state,

billing\_city AS city,

COUNT(invoice\_id) AS number\_of\_invoices,

SUM(total) AS total\_revenue

FROM

invoice

GROUP BY

billing\_country,

billing\_state,

billing\_city

ORDER BY

total\_revenue DESC;

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

**Ans-:**

**Insights:**

* **Top customers differ across countries.**

**Recommendations:**

1. **Launch targeted campaigns to engage top customers in each country effectively.**
2. **Introduce tiered loyalty programs to reward customers based on their spending levels.**
3. **Categorize customers by location and value to enable tailored communication strategies.**
4. **Examine purchase history and preferences of top customers to identify factors driving their high value and refine strategies accordingly.**

My Sql Query-: WITH CustomerRevenue AS (

SELECT

c.customer\_id,

c.first\_name,

c.last\_name,

c.country,

SUM(i.total) AS total\_revenue

FROM

customer AS c

JOIN

invoice AS i ON c.customer\_id = i.customer\_id

GROUP BY

c.customer\_id, c.country

),

RankedCustomers AS (

SELECT

customer\_id,

first\_name,

last\_name,

country,

total\_revenue,

RANK() OVER (PARTITION BY country ORDER BY total\_revenue DESC) AS revenue\_rank

FROM

CustomerRevenue

)

SELECT

customer\_id,

first\_name,

last\_name,

country,

total\_revenue

FROM

RankedCustomers

WHERE

revenue\_rank <= 5

ORDER BY

country, total\_revenue DESC;

6. Identify the top-selling track for each customer

**Ans-: Identifying a top selling track for each customer**

**is difficult as the distribution of tracks bought by customers**

**doesn’t seem to be differentiating.**

**My Sql Query -: WITH CustomerTrackSales AS (**

**SELECT**

**i.customer\_id,**

**t.track\_id,**

**t.name AS track\_name,**

**SUM(il.quantity) AS total\_quantity\_sold**

**FROM**

**invoice\_line AS il**

**JOIN**

**invoice AS i ON il.invoice\_id = i.invoice\_id**

**JOIN**

**track AS t ON il.track\_id = t.track\_id**

**GROUP BY**

**i.customer\_id, t.track\_id**

**),**

**TopTracks AS (**

**SELECT**

**customer\_id,**

**track\_name,**

**total\_quantity\_sold,**

**RANK() OVER (PARTITION BY customer\_id ORDER BY total\_quantity\_sold DESC) AS track\_rank**

**FROM**

**CustomerTrackSales**

**)**

**SELECT**

**customer\_id,**

**track\_name AS top\_selling\_track,**

**total\_quantity\_sold**

**FROM**

**TopTracks**

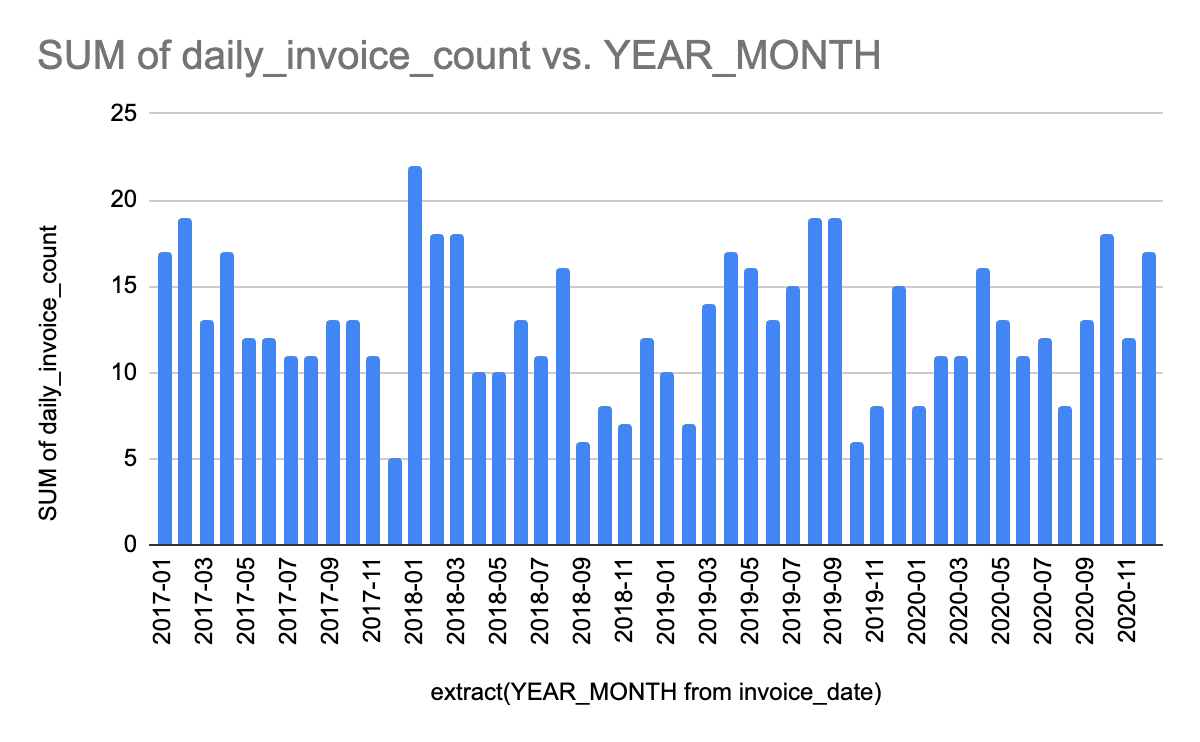
**WHERE**

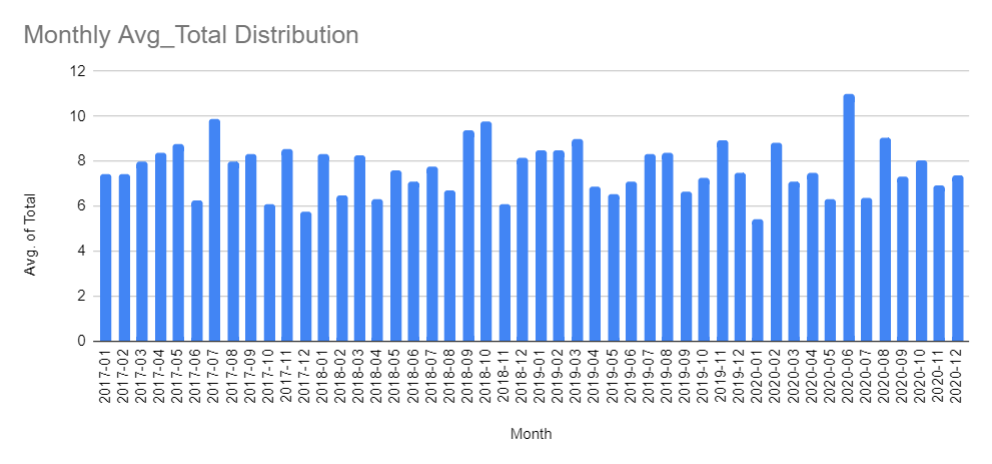
**track\_rank = 1**

**ORDER BY**

**customer\_id;**

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

**Ans-: The given chart summarizes the monthly distribution of Total billing of invoices across the 4 year period and the monthly distribution of Invoices across the 4 year period**.



**Insights:**

1. **Sales show clear peaks in December 2018 and January 2019, indicating increased activity during the holiday season and year-end.**
2. **Monthly invoice counts are generally consistent, though there are dips mid-year (around June and July) and increases toward year-end, reflecting mild seasonality aligned with overall sales trends.**
3. **While monthly averages hover around 7–8 units, occasional spikes above 9 (e.g., mid-2018 and early 2020) suggest occurrences of high-value transactions during these periods.**

**Recommendations:**

1. **Maximize revenue during peak months (December and January) by implementing targeted marketing campaigns, promotions, and product bundling to attract more customers and increase spending.**
2. **Counteract the sales decline in June and July by launching mid-year promotions or loyalty programs. Capitalize on the slight uplift in late April and mid-May to sustain momentum.**
3. **Analyze customer behavior and strategies from high-performance months (e.g., mid-2018, early 2020) to identify factors driving high-value transactions and replicate these tactics in other months.**

**My SQL Query-: SELECT**

**c.customer\_id,**

**c.first\_name,**

**c.last\_name,**

**c.country,**

**COUNT(i.invoice\_id) AS number\_of\_purchases**

**FROM**

**customer AS c**

**JOIN**

**invoice AS i ON c.customer\_id = i.customer\_id**

**GROUP BY**

**c.customer\_id**

**ORDER BY**

**number\_of\_purchases DESC; -- number of purchases by customer**

**select customer\_id, avg(total) as avg\_order\_value, count(invoice\_id)as num\_of\_orders**

**from invoice**

**group by customer\_id**

**order by count(invoice\_id),avg(total); -- avg.order values and number of order**

**select count(invoice\_id) as daily\_invoice\_count, extract(YEAR\_MONTH from invoice\_date), avg(total) as monthly\_avg\_total, sum(total) as monthly\_sum\_total**

**from invoice**

**group by extract(YEAR\_MONTH from invoice\_date)**

**order by extract(YEAR\_MONTH from invoice\_date);**

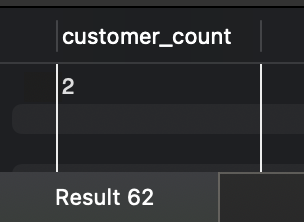
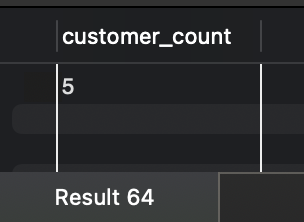
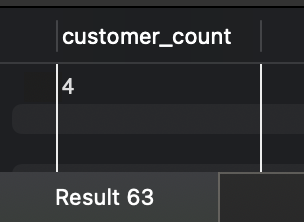
**-- daily invoices and monthly sum totals**

8. What is the customer churn rate?

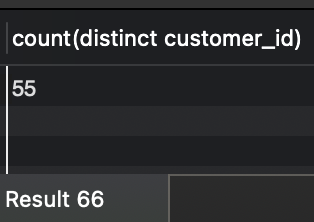
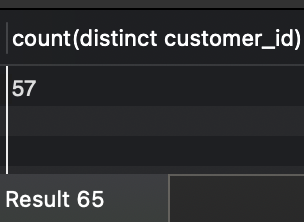
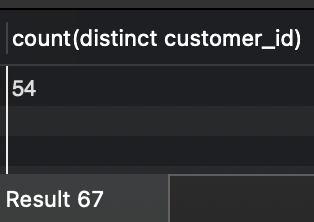
**Ans-: The formula of calculating the churn rate is**

**Churn rate = Total Churned Customers\*100​/Average Customer Base.**

**And for calculating the total churned customer we have to add the all churned customer of every year.**

**Total churned customer = churned customers of 2018 + churned customers of 2019 + churned customers of 2020 (2+4+5) = 11** 

**Avg customer base = (customer in 2018 + customer in 2019 + customer in 2020) / 3**

**(55+57+54)/3 = 55.3**

**Churn rate = 11\*100/55.3 = 19.8%**

**Insights:**

1. **High Churn Rate (19.8%): Nearly 20% of customers are lost annually, signaling a significant churn issue.**
2. **While overall customer numbers remain steady, retention efforts appear to be declining.**

**Recommendations:**

1. **Implement loyalty programs and personalized offers to encourage repeat business and strengthen customer loyalty.**
2. **Gather customer feedback to identify pain points and address the reasons why customers leave.**
3. **Focus on engaging new customers early with clear communication, support, and incentives to build lasting relationships.**

**My SQL Query-: -- Total Number of Customers**

**SELECT COUNT(DISTINCT customer\_id) AS total\_customers**

**FROM customer;**

**-- Number of Active Customers**

**select \* from invoice;**

**SELECT COUNT(DISTINCT customer\_id) AS active\_customers**

**FROM invoice**

**WHERE DATE(invoice\_date) >= '2020-12-30';**

**-- Number of Churned Customers**

**SELECT COUNT(DISTINCT c.customer\_id) AS churned\_customers**

**FROM customer c**

**WHERE c.customer\_id NOT IN (**

**SELECT DISTINCT customer\_id**

**FROM invoice**

**WHERE DATE(invoice\_date) >= '2020-12-30'**

**);**

**-- now count the distinct customer by every year and churn customer by every year**

**select count(distinct customer\_id) as customer\_count from invoice**

**where invoice\_date between '2018-01-01' and '2018-12-31' and customer\_id not in**

**(select distinct customer\_id from invoice**

**where invoice\_date between '2017-01-01' and '2017-12-31');**

**-- customers churned in 2018**

**select count(distinct customer\_id) as customer\_count from invoice**

**where invoice\_date between '2019-01-01' and '2019-12-31' and customer\_id not in**

**(select distinct customer\_id from invoice**

**where invoice\_date between '2018-01-01' and '2018-12-31');**

**-- customers churned in 2019**

**select count(distinct customer\_id) as customer\_count from invoice**

**where invoice\_date between '2020-01-01' and '2020-12-31' and customer\_id not in**

**(select distinct customer\_id from invoice**

**where invoice\_date between '2019-01-01' and '2019-12-31');**

**-- customers churned in 2020**

**select count(distinct customer\_id) from invoice**

**where invoice\_date between '2017-01-01' and '2017-12-31';**

**-- customers in 2018**

**select count(distinct customer\_id) from invoice**

**where invoice\_date between '2018-01-01' and '2018-12-31';**

**-- customers in 2019**

**select count(distinct customer\_id) from invoice**

**where invoice\_date between '2019-01-01' and '2019-12-31';**

**-- customers in 2020**

**with first\_three\_months as**

**(**

**select count(customer\_id) as customer\_count from invoice**

**where invoice\_date between '2017-01-01' and '2017-03-31'**

**),**

**last\_three\_months as**

**(**

**select count(customer\_id) as customer\_count from invoice**

**where invoice\_date between '2017-10-01' and '2017-12-31'**

**)**

**select ((first3.customer\_count)-(last3.customer\_count))/(first3.customer\_count) \* 100 as churn\_rate**

**from first\_three\_months as first3,last\_three\_months as last3;**

**-- Churn Rate in 2017**

**with first\_three\_months as**

**(**

**select count(customer\_id) as customer\_count from invoice**

**where invoice\_date between '2018-01-01' and '2018-03-31'**

**),**

**last\_three\_months as**

**(**

**select count(customer\_id) as customer\_count from invoice**

**where invoice\_date between '2018-10-01' and '2018-12-31'**

**)**

**select ((first3.customer\_count)-(last3.customer\_count))/(first3.customer\_count) \* 100 as churn\_rate**

**from first\_three\_months as first3,last\_three\_months as last3;**

**-- churn rate in 2018**

**with first\_three\_months as**

**(**

**select count(customer\_id) as customer\_count from invoice**

**where invoice\_date between '2019-01-01' and '2019-03-31'**

**),**

**last\_three\_months as**

**(**

**select count(customer\_id) as customer\_count from invoice**

**where invoice\_date between '2019-10-01' and '2019-12-31'**

**)**

**select ((first3.customer\_count)-(last3.customer\_count))/(first3.customer\_count) \* 100 as churn\_rate**

**from first\_three\_months as first3,last\_three\_months as last3;**

**-- churn rate in 2019**

**with first\_three\_months as**

**(**

**select count(customer\_id) as customer\_count from invoice**

**where invoice\_date between '2020-01-01' and '2020-03-31'**

**),**

**last\_three\_months as**

**(**

**select count(customer\_id) as customer\_count from invoice**

**where invoice\_date between '2020-10-01' and '2020-12-31'**

**)**

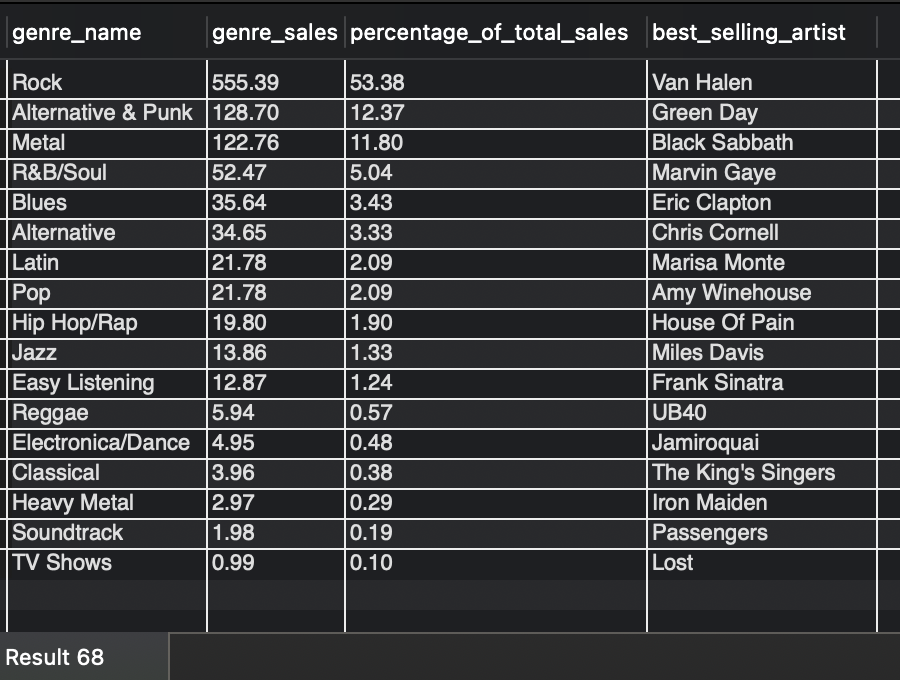
**select ((first3.customer\_count)-(last3.customer\_count))/(first3.customer\_count) \* 100 as churn\_rate**

**from first\_three\_months as first3,last\_three\_months as last3;**

**-- churn rate in 2020.**

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

**Ans-: The best-selling genre in the USA is Rock, contributing 53.38% of the total revenue.**

**Despite this, the best-selling artist is Van Halen, indicating their appeal transcends the dominance of any single genre.(That's what we can see in our Obj. que-2)**

**Insights:**

1. **Rock music contributes over 50% of total sales in the US, solidifying its status as the leading genre.**
2. **Van Halen popularity drives significant sales, making them the best-selling artist.**

**Recommendations:**

1. **Prioritize Rock music in inventory planning and promotional campaigns to maximize revenue.**
2. **Highlight Van Halen in marketing efforts, playlists, and recommendations to leverage their strong customer appeal.**
3. **Invest in promoting genres like Alternative & Punk and Metal, which also show strong sales performance.**
4. **While capitalizing on Rock's popularity, gradually expand the audience by showcasing and marketing other genres to appeal to a broader customer base.**

**My Sql Query-: select \* from genre;**

**SELECT**

**g.name AS genre\_name,**

**i.billing\_state AS region\_state,**

**SUM(il.unit\_price \* il.quantity) AS genre\_sales**

**FROM invoice i**

**JOIN invoice\_line il ON i.invoice\_id = il.invoice\_id**

**JOIN track t ON il.track\_id = t.track\_id**

**JOIN genre g ON t.genre\_id = g.genre\_id**

**WHERE i.billing\_country = 'USA'**

**GROUP BY g.name, i.billing\_state**

**ORDER BY region\_state, genre\_sales DESC; -- by this qurey i check all genre sales in USA.**

**-- now i am Calculating the percentage of total sales contributed by each genre in the USA**

**-- and identify the best-selling genres and artists.**

**WITH GenreSalesUSA AS (**

**SELECT**

**g.name AS genre\_name,**

**SUM(il.unit\_price \* il.quantity) AS genre\_sales**

**FROM invoice i**

**JOIN invoice\_line il ON i.invoice\_id = il.invoice\_id**

**JOIN track t ON il.track\_id = t.track\_id**

**JOIN genre g ON t.genre\_id = g.genre\_id**

**WHERE i.billing\_country = 'USA'**

**GROUP BY g.name**

**),**

**TotalSalesUSA AS (**

**SELECT**

**SUM(il.unit\_price \* il.quantity) AS total\_sales**

**FROM invoice i**

**JOIN invoice\_line il ON i.invoice\_id = il.invoice\_id**

**WHERE i.billing\_country = 'USA'**

**),**

**ArtistGenreSales AS (**

**SELECT**

**g.name AS genre\_name,**

**ar.name AS artist\_name,**

**SUM(il.unit\_price \* il.quantity) AS artist\_genre\_sales,**

**ROW\_NUMBER() OVER (PARTITION BY g.genre\_id ORDER BY SUM(il.unit\_price \* il.quantity) DESC) AS artist\_rank**

**FROM invoice i**

**JOIN invoice\_line il ON i.invoice\_id = il.invoice\_id**

**JOIN track t ON il.track\_id = t.track\_id**

**JOIN album al ON t.album\_id = al.album\_id**

**JOIN artist ar ON al.artist\_id = ar.artist\_id**

**JOIN genre g ON t.genre\_id = g.genre\_id**

**WHERE i.billing\_country = 'USA'**

**GROUP BY g.name, ar.name, g.genre\_id**

**)**

**SELECT**

**gs.genre\_name,**

**gs.genre\_sales,**

**ROUND((gs.genre\_sales / ts.total\_sales) \* 100, 2) AS percentage\_of\_total\_sales,**

**ags.artist\_name AS best\_selling\_artist**

**FROM GenreSalesUSA gs**

**JOIN TotalSalesUSA ts ON 1=1**

**LEFT JOIN ArtistGenreSales ags ON gs.genre\_name = ags.genre\_name AND ags.artist\_rank = 1**

**ORDER BY gs.genre\_sales DESC;**

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

**Ans-: All 59 customers have purchased records spanning at least three different genres, showcasing diverse musical preferences.**

**Leonie Kohler stands out as the top buyer among them, reflecting her significant contribution to sales.**

**Insights:**

1. **All customers have bought records from at least three different genres, indicating broad musical interests.**
2. **Leonie Köhler leads as the top buyer, purchasing from 14 different genres.**

**Recommendations:**

1. **Use customer purchase data to recommend tracks in genres they haven’t yet explored, expanding their engagement.**
2. **Design promotions tailored to customers’ specific genre preferences to drive repeat purchases.**
3. **Introduce a loyalty program that rewards customers for exploring and purchasing from new genres, such as discounts on concert tickets.**
4. **Study overlaps between genres to identify synergistic opportunities for bundling or cross-promotion.**

**My SQL Query-:**

**SELECT**

**c.customer\_id,**

**c.first\_name,**

**c.last\_name,**

**COUNT(DISTINCT g.genre\_id) AS genre\_count**

**FROM customer c**

**JOIN invoice i ON c.customer\_id = i.customer\_id**

**JOIN invoice\_line il ON i.invoice\_id = il.invoice\_id**

**JOIN track t ON il.track\_id = t.track\_id**

**JOIN genre g ON t.genre\_id = g.genre\_id**

**GROUP BY c.customer\_id, c.first\_name, c.last\_name**

**HAVING COUNT(DISTINCT g.genre\_id) >= 3**

**ORDER BY genre\_count DESC;**

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

**Ans-: Rock is the top-performing genre by a significant margin.**

**My SQL Query-:**

**SELECT**

**g.name AS genre\_name,**

**SUM(il.unit\_price \* il.quantity) AS total\_sales,**

**RANK() OVER (ORDER BY SUM(il.unit\_price \* il.quantity) DESC) AS genre\_rank**

**FROM invoice i**

**JOIN invoice\_line il ON i.invoice\_id = il.invoice\_id**

**JOIN track t ON il.track\_id = t.track\_id**

**JOIN genre g ON t.genre\_id = g.genre\_id**

**WHERE i.billing\_country = 'USA'**

**GROUP BY g.genre\_id, g.name**

**ORDER BY genre\_rank;**

**Insights:**

1. **Rock is the top-performing genre by a significant margin.**
2. **The top three genres—Rock, Alternative & Punk, and Metal—highlight strong customer preference for guitar-centric styles.**

**Recommendations:**

1. **Ensure a balanced stock of Rock, Alternative & Punk, and Metal to meet demand efficiently.**
2. **Concentrate marketing efforts on Rock and related genres to capitalize on existing preferences.**
3. **Run targeted campaigns like "Rocktober" sales to boost engagement and sales during specific periods.**
4. **Investigate customer preferences within subgenres (e.g., classic rock, hard rock) and tailor promotions to these niches.**
5. **Develop strategies to introduce and market less popular genres, expanding the customer base and fostering broader engagement over time.**

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

**Ans-: These are the customers who didn't make any purchases in the last three months.**

**My Sql Query-:**

**select \* from invoice;**

**select first\_name, last\_name from customer c**

**left join (**

**select \***

**from invoice**

**where invoice\_date >**

**(select max(invoice\_date) from invoice)**

**- interval 3 month) prev\_3\_months**

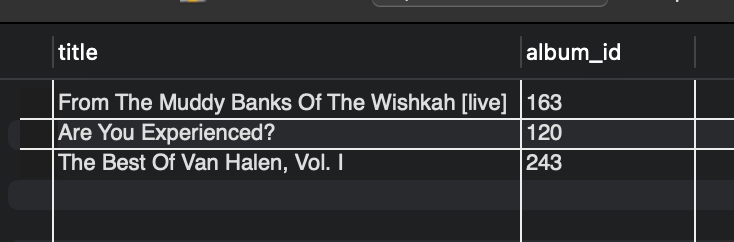
**on prev\_3\_months.customer\_id = c.customer\_id**

**where invoice\_id is null;**

Subjective Questions

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

**Ans-: Based on the analysis in Question 11, Rock is the most popular genre in the USA according to sales data. Consequently, the top three albums from the Rock genre should be prioritized for advertising and promotion efforts in the USA.**

**Here are the top three albums ranked by the quantity of tracks sold:**

**My SQL Query -:**

**select \* from track order by album\_id,genre\_id; -- an album have songs with same/single genre**

**select genre\_id from genre where name = 'Rock'; -- genre\_id = 1 for Rock**

**with cte as(**

**select sum(i.total) as total\_revenue, t.album\_id**

**from invoice i left join invoice\_line il on il.invoice\_id = i.invoice\_id**

**left join track t on t.track\_id = il.track\_id**

**where i.billing\_country = 'USA' and t.genre\_id = 1**

**group by t.album\_id**

**order by total\_revenue desc**

**)**

**select a.title, a.album\_id from album a**

**left join cte on cte.album\_id = a.album\_id**

**order by cte.total\_revenue desc limit 3;**

**Insights:**

1. **Rock remains the most preferred genre in the USA.**
2. **The top three Rock albums are:**
   * ***From The Muddy Banks Of The Wishkah [Live]***
   * ***Are You Experienced?***
   * ***The Best Of Van Halen, Vol. I***

**Recommendations:**

1. **Concentrate advertising and promotional efforts on these three albums in the US market to maximize impact.**
2. **Introduce special bundles or discounts combining these albums to encourage higher sales.**
3. **Leverage data to identify US customers who have purchased Rock music and target them with tailored advertisements for these albums.**

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

**Ans-: Based on the data, it is evident that Rock is the unifying preference between the USA and the rest of the world, as it holds the top position in both cases. However, beyond this shared preference for Rock, there appear to be no other significant commonalities between the two datasets.**

**Other countries USA.**

**My SQL Query-: SELECT Top\_Genre FROM**

**(**

**select g.name as Top\_Genre**

**from track t**

**left join invoice\_line il on il.track\_id = t.track\_id**

**left join invoice i on i.invoice\_id = il.invoice\_id**

**left join genre g on t.genre\_id = g.genre\_id**

**where i.billing\_country != 'USA'**

**group by g.name**

**order by sum(il.quantity) desc**

**limit 10**

**) sub\_table;**

**Insights:**

1. **Rock maintains its dominance as the top genre both in the USA and globally.**
2. **Beyond Rock, the USA's genre preferences differ significantly from other countries.**
3. **Genres like Metal, Latin, and Jazz exhibit stronger performance outside the USA.**

**Recommendations:**

1. **Maintain a strong focus on Rock music across all markets to capitalize on its widespread popularity.**
2. **Tailor inventory and promotional efforts to align with local preferences, such as Metal, Latin, and Jazz outside the USA.**
3. **Design region-specific marketing campaigns that resonate with the tastes and trends of local audiences.**
4. **Strategize ways to introduce and promote less popular genres in both US and international markets to expand the customer base.**

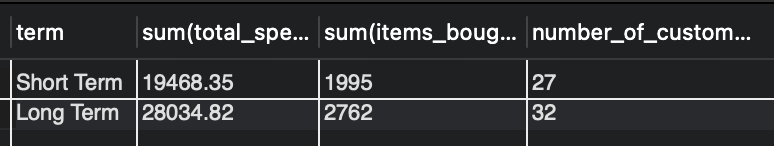
3. Customer Purchasing Behavior Analysis: How do the purchasing habits (frequency, basket size, spending amount) of long-term customers differ from those of new customers? What insights can these patterns provide about customer loyalty and retention strategies?

**Ans-: Average Association Period Calculation: Determines the average customer association period with Chinook.**

**Customer Segmentation:**

* **Classified customers as Long-Term if their association period > average.**
* **Classified customers as Short-Term if their association period ≤ average.**

**Metric Analysis:**

* **Calculated key metrics for each group, including:**
  + **Total expenditure**
  + **Items purchased**
  + **Purchase frequency**

**Recommendations:**

1. **Enhance Customer Loyalty:**
   * **Recognize that customer loyalty is critical for long-term revenue generation at Chinook.**
   * **Focus on retaining long-term customers while encouraging short-term customers to stay longer.**
2. **Offer Deals and Discounts:**
   * **Introduce loyalty programs, exclusive deals, and periodic discounts tailored for long-term customers.**
   * **Provide special offers to short-term customers to encourage repeat purchases and extend their association period.**
3. **Prioritize User Experience:**
   * **Ensure seamless and enjoyable user experiences across all touchpoints, from browsing to purchase and post-sale services.**
   * **Gather customer feedback regularly to identify areas of improvement and address pain points promptly.**
4. **Personalized Engagement:**
   * **Use data-driven insights to deliver personalized recommendations, promotions, and communications based on customer preferences and behavior.**

**My SQL Query -: WITH cte as**

**(**

**select i.customer\_id, max(invoice\_date) as last\_purchase\_date, min(invoice\_date) as first\_purchase\_date,**

**sum(total) as total\_spent, sum(quantity) as items\_bought, count(i.customer\_id) as frequency,**

**abs(timestampdiff(day, max(invoice\_date), min(invoice\_date))) as customer\_since\_days**

**from invoice i**

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

**left join customer c on c.customer\_id = i.customer\_id**

**group by i.customer\_id**

**),**

**long\_short\_term as**

**(**

**SELECT total\_spent, items\_bought, frequency,**

**case**

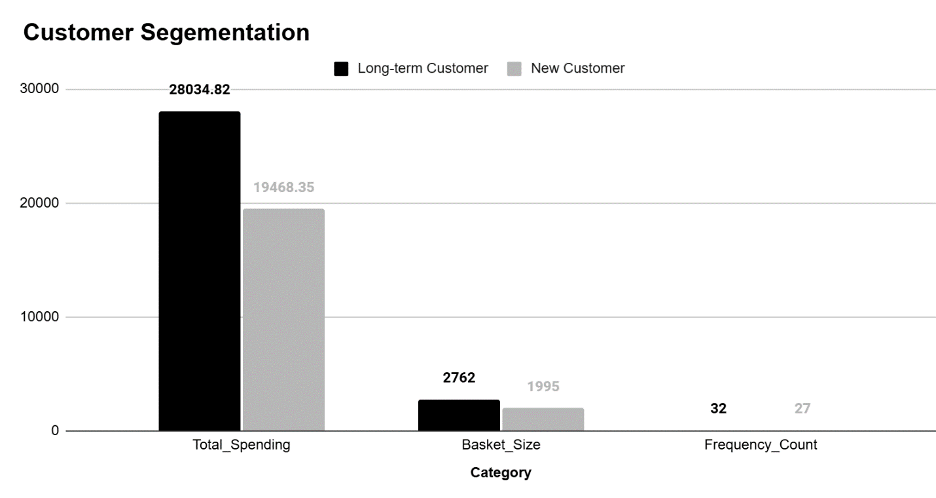
**when customer\_since\_days>(select avg(customer\_since\_days) as average\_days from cte) then 'Long Term'**

**else 'Short Term' end term**

**from cte**

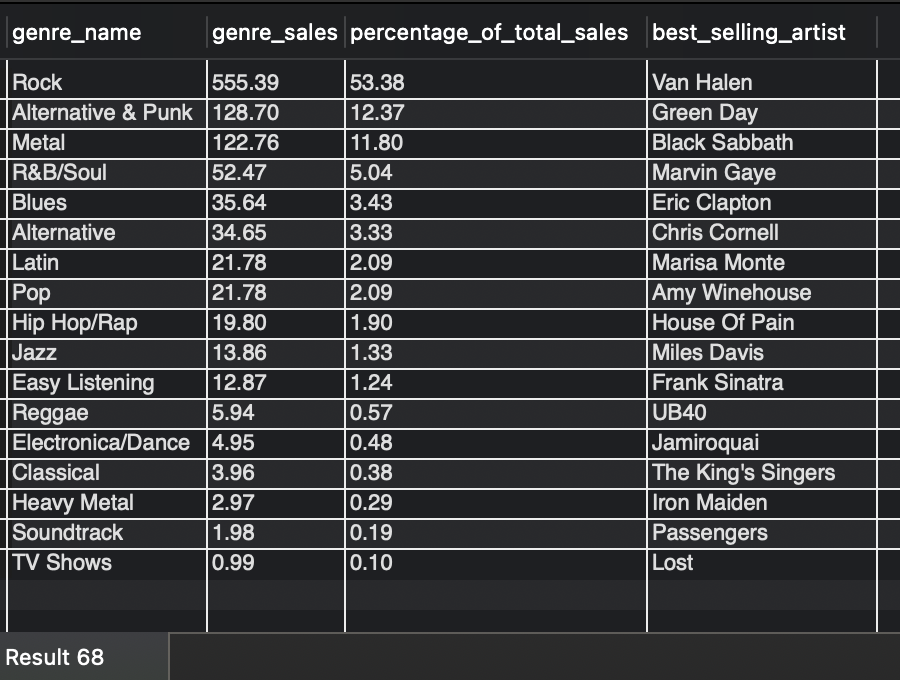
**)**

**select term, sum(total\_spent),sum(items\_bought),count(frequency) as number\_of\_customers from long\_short\_term group by term;**

**Visualization -: **

4. 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?

**Ans-:**



**My SQL Query-:**

**select \* from invoice\_line;**

**select il.invoice\_id,g.name**

**from invoice\_line il**

**left join track t on t.track\_id = il.track\_id**

**left join genre g on g.genre\_id = t.genre\_id**

**group by il.invoice\_id,g.name;**

**-- different genres purchased over an invoice**

**select il.invoice\_id, al.title**

**from invoice\_line il**

**left join track t on t.track\_id = il.track\_id**

**left join album al on al.album\_id = t.album\_id**

**group by il.invoice\_id, al.title;**

**-- different albums purchased over an invoice**

**select il.invoice\_id,a.name**

**from invoice\_line il**

**left join track t on t.track\_id = il.track\_id**

**left join album al on al.album\_id = t.album\_id**

**left join artist a on a.artist\_id = al.artist\_id**

**group by il.invoice\_id,a.name ;**

**-- different artists preferred in a single invoice**

**Insights:**

1. **Most Preferred Genres:**
   * **The pivot table analysis shows that Rock, Metal, and Alternative & Punk are the top three preferred genres by customers.**
   * **These genres have strong overlap, making them ideal for cross-promotion and recommendations.**
2. **Cross-Selling Opportunity:**
   * **If a customer purchases from one of these genres, salespeople can recommend other albums or tracks from the other two genres, capitalizing on their preferences.**

**Recommendations:**

1. **Cross-Promotions:**
   * **Encourage sales staff to recommend albums from Metal and Alternative & Punk to customers buying Rock albums, and vice versa.**
   * **Highlight similarities in style, instrumentation, and overall musical appeal to boost the likelihood of additional purchases.**
2. **Genre Bundles:**
   * **Create bundled offers that combine albums or tracks from Rock, Metal, and Alternative & Punk genres, giving customers the chance to explore related music at a discounted price.**
3. **Personalized Recommendations:**
   * **Implement a recommendation system (either through staff or automated tools) to suggest albums from these top genres based on previous purchases.**
   * **Use customer purchase history and genre preferences to automatically trigger suggestions for related albums.**
4. **Targeted Marketing Campaigns:**
   * **Run marketing campaigns focusing on the overlap of these top genres, highlighting the diversity of offerings within Rock, Metal, and Alternative & Punk for customers who prefer one genre.**
   * **Use this data to segment customers and send personalized email offers or notifications about new releases in these genres.**
5. **Event and Content Integration:**
   * **Use live events, such as concerts or genre-specific playlists, to encourage exploration of similar genres.**
   * **Promote content, such as "If you like Rock, you’ll love Metal and Alternative & Punk," in both online and in-store settings.**

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

**Ans-:Insights:**

1. **Low or Negative Churn in Developed Economies:**
   * **Countries such as the USA, Canada, Germany, UK, Australia, and Brazil have low or negative churn rates.**
   * **This trend might be influenced by their status as developed economies with higher GDP, which provides greater market stability and purchasing power.**
2. **Higher Churn in Developing or Less Developed Economies:**
   * **Countries like Norway, Netherlands, Chile, Denmark, and India show comparatively higher churn rates.**
   * **This could indicate challenges in customer retention and a decreasing business scope in these regions.**

**Recommendation:**

1. **Targeted Advertising in High-Economy Countries:**
   * **Increase the frequency and visibility of advertisements in developed countries with high GDP, such as the USA, Canada, Germany, UK, and Australia.**
   * **Highlight premium offerings and exclusive content in these markets to capitalize on their higher purchasing power.**
2. **Affordable Pricing in Low-Economy Countries:**
   * **Introduce budget-friendly tracks or subscription plans in developing or less developed countries like India, Chile, and Denmark.**
   * **Promote localized discounts and value-driven bundles to attract cost-sensitive customers in these regions.**

**My SQL Query-: with first\_six\_months as**

**(**

**select billing\_country, COUNT(customer\_id) counter from invoice**

**where invoice\_date between '2017-01-01' and '2017-06-30'**

**group by billing\_country**

**),**

**last\_six\_months as**

**(**

**select billing\_country, COUNT(customer\_id) counter from invoice**

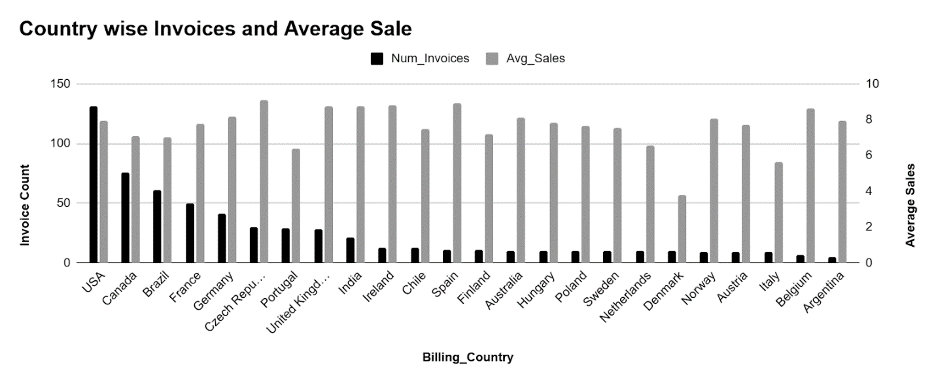
**where invoice\_date between '2020-07-01' and '2020-12-31'**

**group by billing\_country**

**)**

**select f6.billing\_country, (f6.counter - coalesce(l6.counter,0))/f6.counter \* 100 churn\_rate from first\_six\_months f6**

**left join last\_six\_months l6 on f6.billing\_country = l6.billing\_country;**

**Visualization-: **

6. 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?

**Ans-:** **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?**

**My Sql Query- Select i.customer\_id, concat(c.first\_name, " ", c.last\_name) as customer\_name, i.billing\_country,**

**sum(i.total) as total\_spending, COUNT(i.invoice\_id) as num\_of\_orders**

**from invoice i**

**left join customer c on c.customer\_id = i.customer\_id**

**group by i.customer\_id, concat(c.first\_name, " ", c.last\_name), i.billing\_country**

**order by total\_spending desc, num\_of\_orders desc;**

**OUTPUT -**

**Visualization-**

**After analysing the data it can be seen that the countries with already high spending amount and frequency of orders, their numbers are increasing whereas the sales and frequency are stagnant in other countries. Therefore, it is seen that new promotional campaigns need to be done in those countries to reduce churn rate as well as maintain & increase the spending.**

**If information given about age and gender following factors contribute to risk are: -**

**· Are younger customers more likely to churn?**

**· Does gender play a role?**

**· How does location impact churn?**

**· Analysis of spending behaviour**

**If there was information regarding age and gender of the customers, the customer segmentation would have been –**

**· Young-Male-With-High-Spending**

**· Young-Female-With-High-Spending**

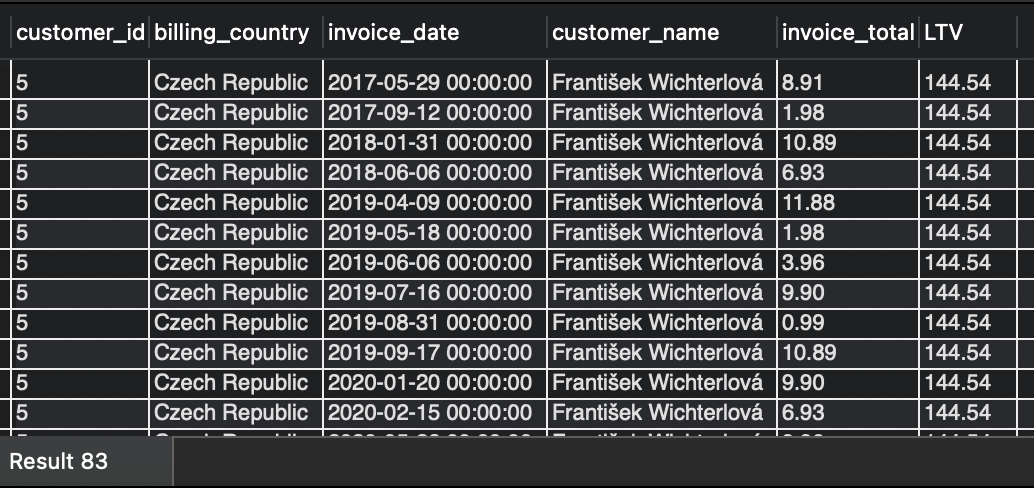
**· Old-Male-With-Low-Spending**

**Old-Female-With-Low-Spending**

7. Customer Lifetime Value Modeling: 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?

**Ans-:Insights:**

1. **Customer Segmentation Potential:**
   * **Data like purchase history, tenure, and engagement can reveal valuable insights into customer traits, such as loyalty, value, and buying frequency.**
   * **Recent high-value customers have the potential to become long-term loyal customers.**
2. **Targeted Marketing Opportunities:**
   * **Customers segmented by purchase behavior can be targeted with personalized campaigns to drive engagement and sales.**
3. **Regional Churn Disparities:**
   * **Higher churn rates in less developed countries might be driven by issues like poor product-market fit or pricing mismatches.**



**Recommendations:**

1. **Leverage Data for Segmentation:**
   * **Segment customers based on purchase patterns (e.g., high-value, frequent buyers, new customers).**
   * **Use these segments to design personalized campaigns and offers.**
2. **Focus on Retention of High-Potential Customers:**
   * **Identify recent high-value customers and target them with loyalty programs, exclusive offers, and proactive engagement.**
3. **Address Churn in Less Developed Countries:**
   * **Conduct market research to better understand local needs and preferences.**
   * **Adjust pricing strategies and product offerings to align with these markets.**
4. **Expand Product-Market Fit Analysis:**
   * **Regularly assess how well products align with customer preferences in different regions.**
   * **Refine offerings to suit local demands while maintaining global appeal.**
5. **Deploy Regional Campaigns:**
   * **Tailor promotional efforts based on economic conditions and customer profiles in specific regions to maximize effectiveness.**

**My SQL Query-: with cte as(select inv.customer\_id,inv.billing\_country,inv.invoice\_date, concat(c.first\_name,' ',c.last\_name) as customer\_name,inv.total as invoice\_total**

**from invoice inv**

**left join customer c on c.customer\_id = inv.customer\_id**

**group by customer\_id,2,3,inv.total**

**order by customer\_name),**

**cte2 as(**

**select customer\_id, sum(total) as LTV**

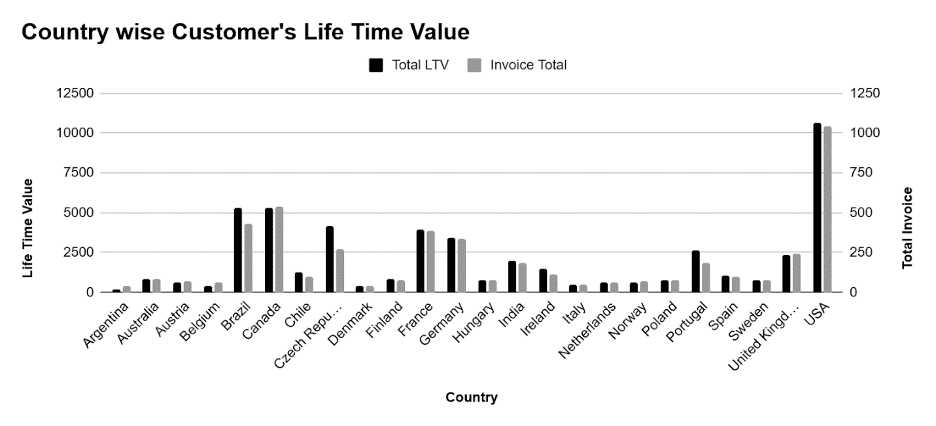
**from invoice**

**group by customer\_id)**

**select cte.customer\_id, cte.billing\_country, cte.invoice\_date, cte.customer\_name,cte.invoice\_total, cte2.LTV**

**from cte left join cte2 on cte.customer\_id = cte2.customer\_id**

**order by cte2.LTV desc,cte.customer\_name, cte.invoice\_date;**

**Visualization-: **

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

**Ans-:Determining which tracks have not had any sales between July 1, 2020, and December 31, 2020.**

**Out of a total of 3503 songs/tracks, 3006 were identified as having no sales.**

**Now Identifying customers who have not made any purchases in the last six months.**

**Out of 59 customers, 16 were identified as having made no purchases in the last six months.**

**Recommendations:**

1. **Targeted Promotions for Tracks: To promote the sales of tracks that have not been sold in the past six months, targeted promotions could be introduced. This could involve offering discounts, bundle deals, or highlighting these tracks in marketing campaigns to generate interest.**
2. **Re-engagement Offers for Inactive Customers: For customers who have been inactive for the last six months, consider launching enticing offers such as special discounts, exclusive deals, or personalized promotions. This could help re-engage these customers and boost business revenue by encouraging repeat purchases.**

**If Data on Promotional Campaigns is Available, We Could Analyze:**

1. **Customer Acquisition:**
   * **Calculate Customer Acquisition Cost (CAC):  
     By calculating the total expenditure on a particular campaign (such as email marketing, events, or discounts) divided by the number of new customers acquired through that campaign, we can identify which campaigns provide the best return on investment by acquiring the most new customers at the lowest cost.**
   * **Attribute New Customers to Specific Campaigns:  
     Ensure that each campaign is linked with a unique code, landing page, or coupon. This would allow us to track where new customers are coming from, giving us a clear understanding of the impact of each campaign.**
2. **Customer Retention:**
   * **Measure Changes in Customer Churn Rate:  
     By tracking churn rates before, during, and after promotional campaigns, we can assess whether specific campaigns help reduce customer churn. For instance, a franchise drive or a product promotion could lead to longer customer retention or conversely, a campaign might cause churn to increase.**
   * **Track Repeat Purchase Rates:  
     We can compare the frequency of repeat purchases between customers who engaged with a promotion and those who did not. This will help evaluate whether promotions foster long-term customer relationships and repeat business.**
3. **Overall Sales:**
   * **Analyze Sales Data During Promotional Periods:  
     Comparing sales figures during promotional campaigns with those from non-promotional periods will help measure the impact of the campaigns on overall sales revenue. This analysis can indicate whether the promotion successfully drove sales growth or had minimal effect.**
   * **Monitor Average Order Value (AOV):  
     By tracking the AOV during promotional periods, we can assess whether promotions, such as discounts or events, encourage customers to spend more per visit, providing insights into the effectiveness of such campaigns in driving larger purchases.**

**My SQL Query-: select count(\*) from track; -- counting total tracks available**

**select t.name**

**from track t**

**where t.track\_id not in ( select il.track\_id**

**from invoice\_line il**

**left join invoice i on i.invoice\_id = il.invoice\_id**

**where i.invoice\_date>='2020-07-01' and i.invoice\_date<='2020-12-31');**

**select concat(c.first\_name,' ',c.last\_name) as full\_name**

**from customer c**

**where c.customer\_id not in (select distinct(customer\_id)**

**from invoice**

**where invoice\_date>='2020-07-01' and invoice\_date<='2020-12-31');**

**-- Identifying customers that have not made any purchase in the previous 6 months.**

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

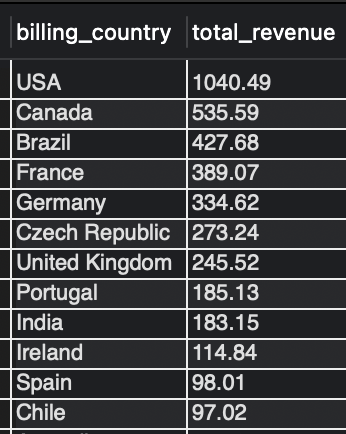
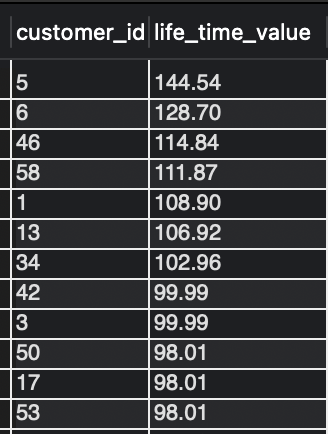
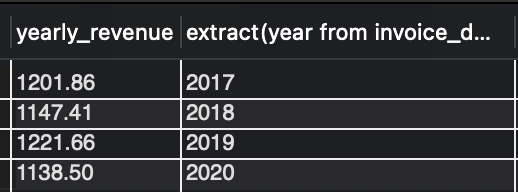
**Ans-:Here's how I'd approach the Chinook music sales data analysis if I wasn't given specific objectives:**

### **1. Data Understanding**

* **Schema Familiarization:  
  Begin by evaluating the schema to understand the structure of the database. This includes identifying the tables, their columns, data types, and the relationships among them.**
* **Data Exploration:  
  Use SQL queries to explore the data. This exploration will involve:**
  + **Checking for Missing Data: Identifying columns with NULL values or incomplete records.**
  + **Identifying Redundancies: Detecting any duplicated records or unnecessary repeated data that could affect analysis**

### **2. Ask Questions Pertaining to the Data**

**After the initial exploration of the data, I would formulate broad, exploratory questions to guide the analysis:**

* **Sales Performance:**
  + **Which albums or tracks have the highest and lowest sales?**
  + **Are there any artists that consistently perform better than others in terms of sales?**
  + **What time of year sees the highest volume of sales? (Can we analyze this by day, week, or month?)**
* **Customer Behavior:**
  + **Who are the top customers based on total spend, and what trends can we spot in their purchase behavior?**
  + **How many repeat customers do we have, and what percentage of total sales come from them?**
  + **Do customers tend to buy single tracks, albums, or a mix of both?**
* **Geographic Insights:**
  + **Are there certain countries or cities with higher sales or specific genre preferences?**
  + **Can we see regional trends in terms of album/track popularity?**
* **Product Popularity:**
  + **Which genres are the most popular among customers?**
  + **Are there any specific tracks that generate consistent sales, or are sales more distributed across multiple tracks?**
* **Revenue Insights:**
  + **How does the pricing of albums/tracks correlate with the volume of sales? Are lower-priced items selling more than premium-priced ones?**
  + **What is the average revenue per customer, and how does that vary by region or customer type?**

**Insights and Recommendations**

### **1. Store Management**

* **Optimize Inventory: Regularly monitor best-selling and underperforming products to manage stock efficiently.**
* **Categorize Products: Use sales data to categorize inventory into high-demand and low-demand products.**

### **2. Targeted Marketing Campaigns**

* **Promote Top-Selling Products: Use targeted campaigns for popular artists and genres. Offer exclusive discounts for repeat customers.**
* **Seasonal & Cross-Selling Promotions: Implement time-limited offers, bundle discounts, and cross-sell related products.**
* **Campaign Timing: Align promotions with peak sales periods based on past trends.**

### **3. Customer Segmentation and Personalization**

* **Segment Customers: Target high-value and inactive customers with tailored offers.**
* **Personalized Recommendations: Suggest albums based on customer preferences and purchase history. Implement dynamic pricing based on demand.**

### **4. New Product Strategies**

* **Introduce New Genres: Use market research to introduce trending genres and launch them with targeted campaigns.**
* **Competitive Pricing: Set competitive prices for new products and offer introductory discounts.**

### **5. Fundamental Strategies**

* **Multi-Channel Marketing: Use a mix of digital ads, email campaigns, and partnerships for cohesive marketing.**
* **Customer Feedback: Regularly survey customers and optimize the user experience based on feedback.**

**My SQL Query-: SELECT \* from employee; -- 1 reports\_to value is null for employee\_id = 1**

**SELECT distinct \* FROM employee; -- No duplicates**

**SELECT \* FROM genre;**

**SELECT distinct \* FROM genre; -- No duplicates**

**SELECT \* FROM invoice;**

**SELECT distinct \* FROM invoice; -- No duplicates**

**SELECT \* FROM invoice\_line;**

**SELECT distinct \* FROM invoice\_line; -- No duplicates**

**SELECT \* FROM media\_type;**

**SELECT distinct \* FROM media\_type; -- No duplicates**

**SELECT \* FROM playlist;**

**SELECT distinct \* FROM playlist; -- No duplicates**

**SELECT \* FROM playlist\_track;**

**SELECT distinct \* FROM playlist\_track;**

**SELECT \* FROM track;**

**SELECT distinct \* FROM track; -- No duplicates**

**select \* from album;**

**select distinct \* from album; -- No duplicates**

**SELECT \* FROM artist;**

**SELECT distinct \* FROM artist; -- No duplicates**

**SELECT \* from customer;**

**SELECT distinct \* FROM customer; -- No duplicates**

**SELECT COUNT(\*) FROM customer;**

**-- WHERE fax is NULL; ( count = 47)**

**-- WHERE state is NULL;(count = 29)**

**-- WHERE company is NULL; (count = 49)**

**-- 47 fax, 29 state and 49 company values are null in the customer table**

**select sum(total) as yearly\_revenue, extract(year from invoice\_date)**

**from invoice**

**group by extract(year from invoice\_date);**

**-- 1201.86 2017**

**-- 1147.41 2018**

**-- 1221.66 2019**

**-- 1138.50 2020**

**select customer\_id, sum(total) as life\_time\_value from invoice**

**group by customer\_id**

**order by life\_time\_value desc;**

**select billing\_country, sum(total) as total\_revenue from invoice**

**group by billing\_country**

**order by total\_revenue desc;**

10. 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?

**Ans-:**

**This is query for this question-: Alter table Album add ReleaseYear int;**

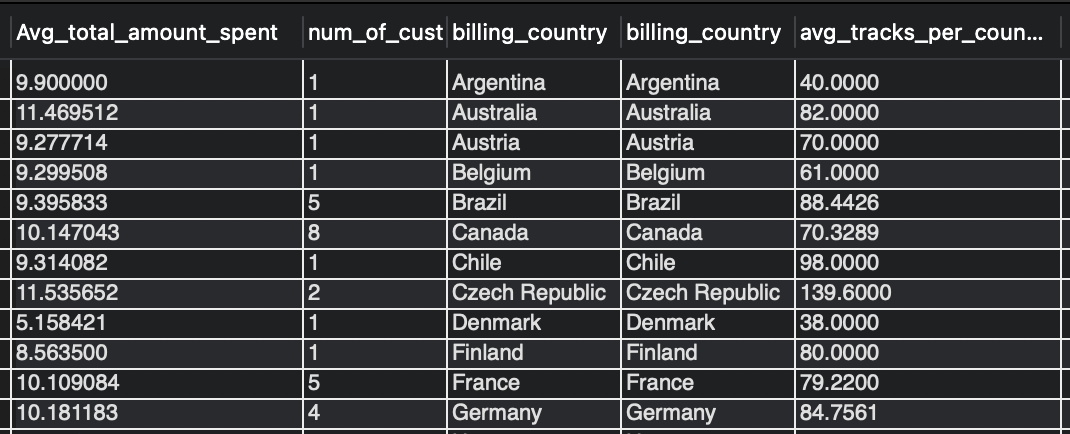
**My SQL Query-:**

**Alter table Album add ReleaseYear int;**

11. 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.

**Ans-:Insights:**

* **Average Spending: The average total amount spent per customer is between 8 to 12, with a notable outlier of 5.158 in Denmark.**
* **Customer Distribution: Most countries have very few customers (single digits), with the USA having the highest number of customers.**

****

### **Recommendations:**

1. **Focus on the USA: Given the larger customer base, prioritize marketing and sales efforts in the USA.**
2. **Increase Penetration in Other Countries: Explore strategies to expand the customer base in countries with fewer customers.**
3. **Tailor Marketing Campaigns: Develop region-specific campaigns based on genre preferences and spending averages to better target customers in different countries.**
4. **Address Low Spending Regions: Analyze low-spending countries and experiment with targeted promotions or product adjustments to boost sales.**
5. **Encourage Higher Spending: Implement loyalty programs and personalized recommendations to incentivize customers to spend more on tracks.**

**My SQL query-: with cte as(**

**select avg(total) Avg\_total\_amount\_spent,**

**count(distinct customer\_id) num\_of\_cust,**

**billing\_country**

**from invoice i**

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

**group by billing\_country**

**),**

**cte2 as(**

**SELECT i.customer\_id, sum(quantity) as quantity\_purchased from invoice i**

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

**group by i.customer\_id**

**),**

**cte3 as(**

**select billing\_country, avg(quantity\_purchased) as avg\_tracks\_per\_country**

**from invoice i**

**left join cte2 on cte2.customer\_id = i.customer\_id**

**group by billing\_country**

**)**

**select \* from cte**

**left join cte3 on cte3.billing\_country = cte.billing\_country;**