Objective Questions

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

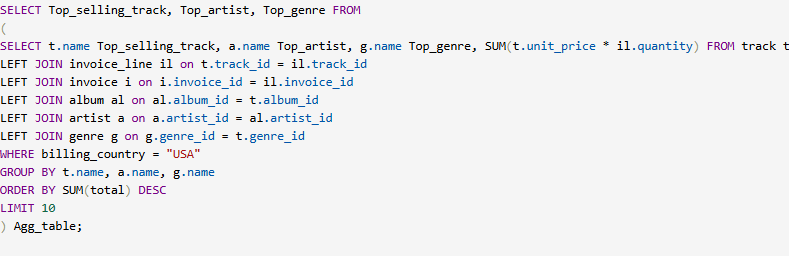
Yes, there were a number of missing values in the dataset like: -

* 49 company, 29 state, 47 fax values are null in the customer table.
* 1 reports\_to value is null in the employee table.
* 978 composer values are null in the track table.

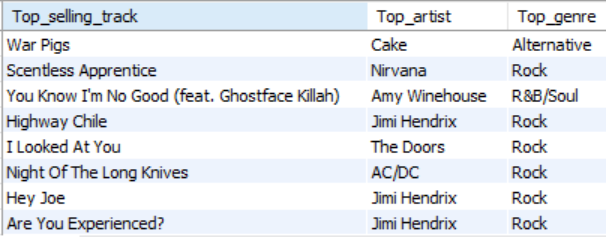
There are no duplicate values in the whole dataset. In case of null values I would use COALESCE function to handle the situation

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

SQL Query: -



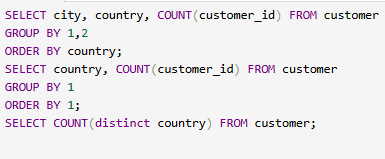
OUTPUT: -



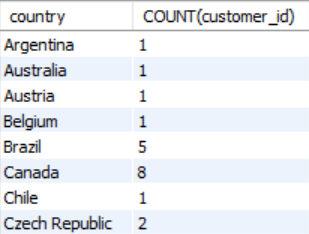
Therefore, the top genre in USA is Rock.

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

SQL Query: -

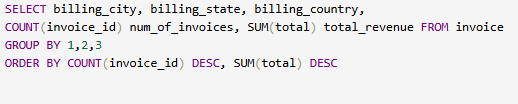


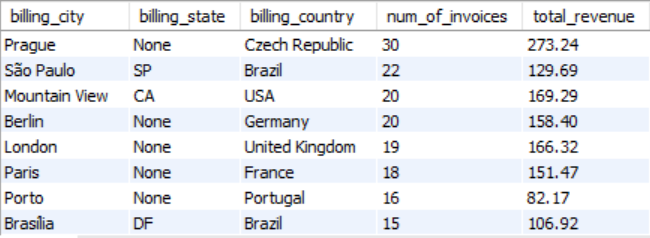
OUTPUT: -

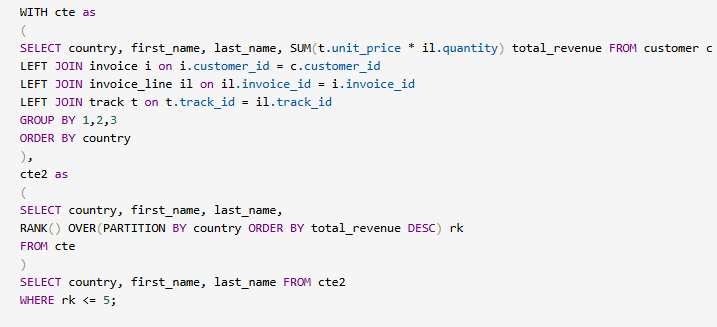
* The Customer Demographic Breakdown based on location is very diversified.
* The Chinook's customer base is spread over 24 countries with max number of customers from USA.
* The customer table does not have age and gender columns to understand the customer breakdown

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

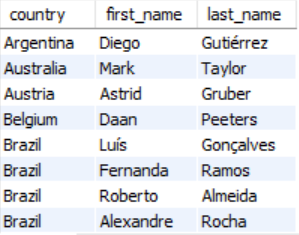
SQL Query: -

OUTPUT: -

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

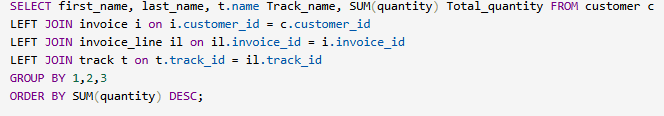
SQL Query: -

OUTPUT: -



1. Identify the top-selling track for each customer

SQL Query: -

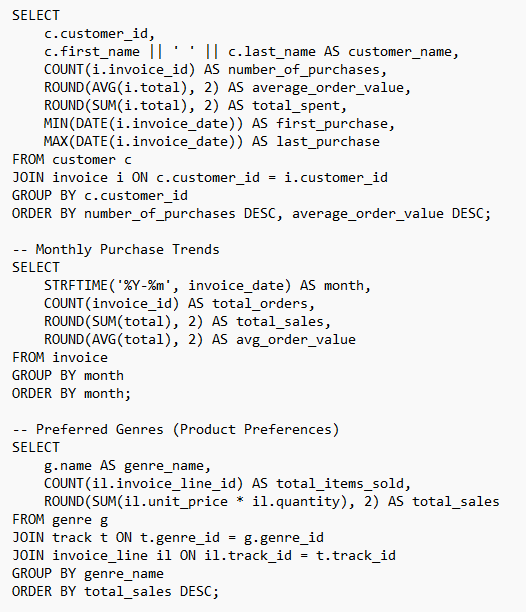
OUTPUT: -



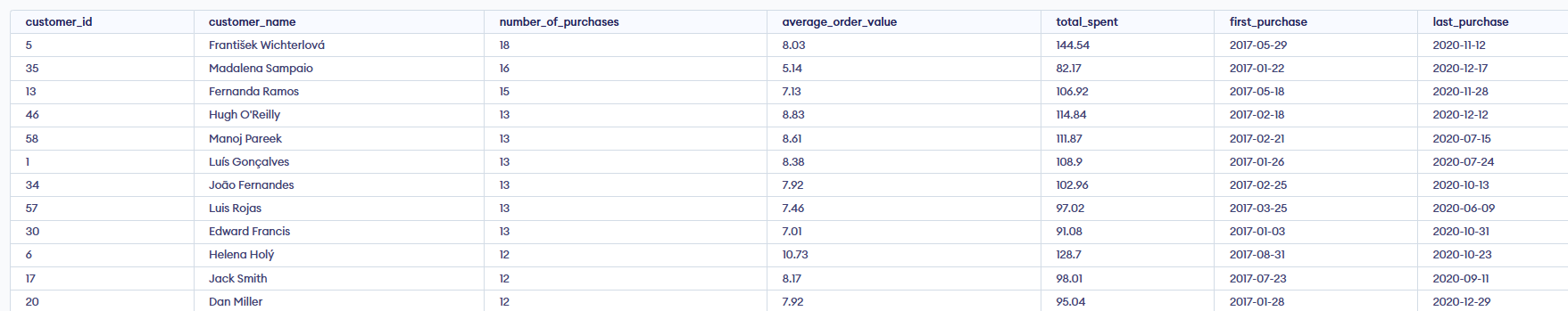
\*The customers had similar quantity of different tracks. So it was difficult to select one as the top track for each customer.

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

SQL Query: -



Output:-



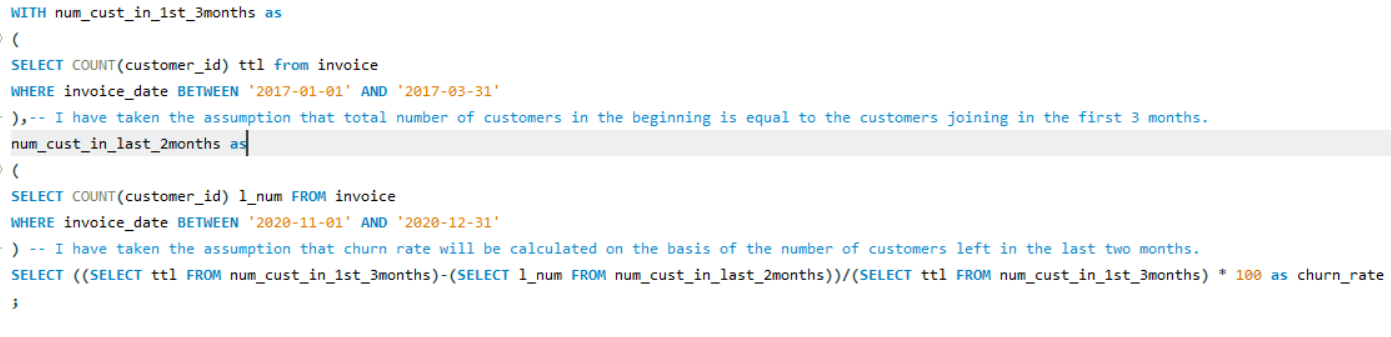
The SQL analysis reveals meaningful patterns in customer behaviour:

* There are **loyal high-spenders** and **casual buyers**.
* Purchase trends **fluctuate monthly**, suggesting seasonal behaviour.
* Certain genres are consistently **more popular**.
* Insights like **average spend** and **purchase frequency** can help **segment customers** and build targeted marketing strategies.

Although payment methods were not available, the rest of the analysis provides strong evidence of customer behaviour patterns using sales and genre data.

1. What is the customer churn rate?

Customer churn rate is defined as the percentage of customers who stop doing business with an organization over a period of time.

SQL Query: -

OUTPUT: -



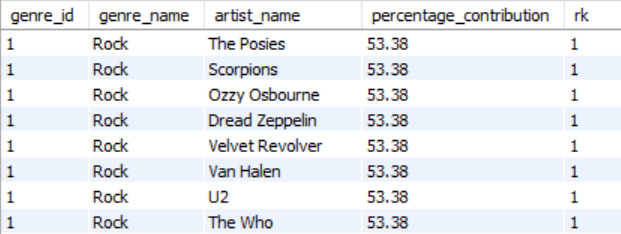
Therefore, the customer churn rate of the company is 40.8163 based on the total number of customers in first 3 months i.e 49 and the number of customers present in the last 2 months i.e 29. So, number of customers lost = 49-29 = 20.

Churn rate = 20/49 = 40.8163

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

SQL Query: -

OUTPUT: -



Therefore, the top selling genre in USA is Rock.

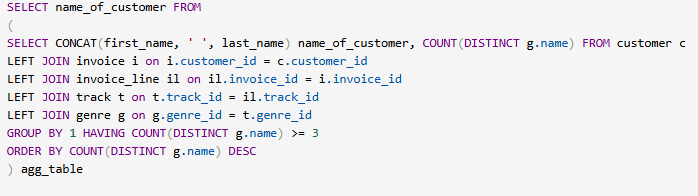
And

The Posies, Scorpions, Ozzy Osbourne, Dread Zeppelin, Velvet Revolver, Van Halen, U2, The Who, The Rolling Stones, The Police, The Doors, The Cult, Terry Bozzio, Tony Levin & Steve Stevens, Stone Temple Pilots, Soundgarden, Skank, Lenny Kravitz, Santana, Rush, Red Hot Chili Peppers, Raul Seixas, R.E.M., Queen, Pink Floyd, Pearl Jam, Paul D'Ianno, Page & Plant, O Terço, Nirvana, Men At Work, Marillion, Led Zeppelin, Kiss, Joe Satriani, Jimi Hendrix, Jamiroquai, Iron Maiden, Guns N' Roses, Foo Fighters, Faith No More, Def Leppard, Deep Purple, Creedence Clearwater Revival, David Coverdale, Frank Zappa & Captain Beefheart, Audioslave, Alice In Chains, Alanis Morissette, Aerosmith, AC/DC, Accept

are all the top artists who are associated with the Rock genre.

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

SQL Query: -

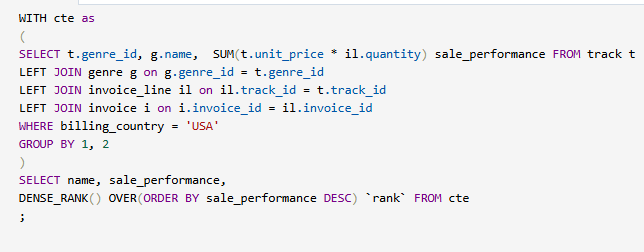
OUTPUT: -



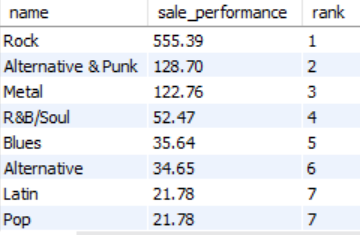
Leonie Köhler is the person who has bought tracks from 14 different genres.

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

SQL Query: -

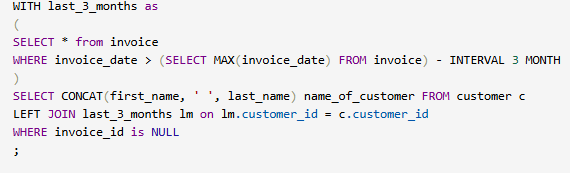


OUTPUT: -

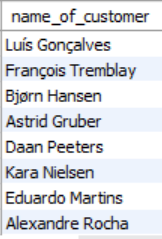


1. Identify customers who have not made a purchase in the last 3 months

SQL Query: -



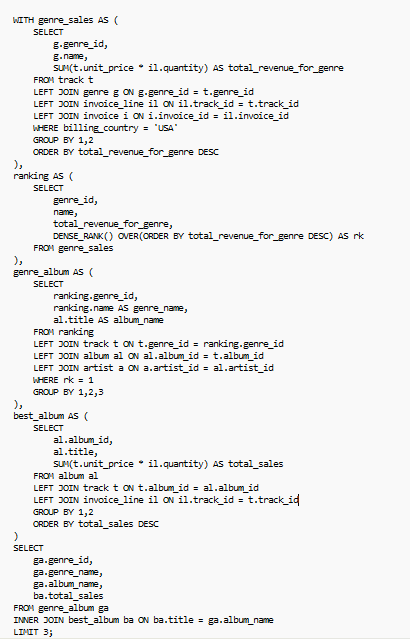
OUTPUT: -



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.

SQL Query: -

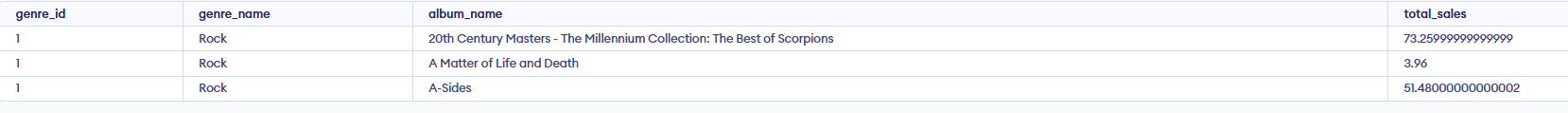


Recommendation: - Top 3 albums that should be prioritised for advertisements and promotions in the USA based on genre analysis are

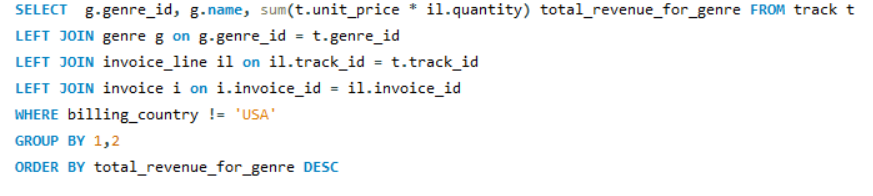
* A Matter of Life and Death
* 20th Century Masters - The Millennium Collection: The Best of Scorpions
* A-Sides

All these albums are from the genre ROCK as it is the most popular genre in the USA.

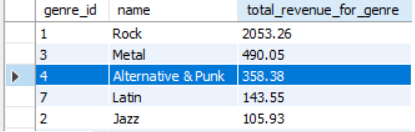
OUTPUT:-



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

SQL Query: - 

Output:-



## **Insight**

### 1. Rock Dominates Globally:

* Rock is by far the top-selling genre outside the USA, with revenue 4x higher than Metal and 6x more than Alternative.
* This indicates universal appeal and strong global fan base.

### 2. Metal & Alternative Are Niche but Profitable:

* Both Metal and Alternative & Punk are consistently purchased genres, showing a loyal international audience.
* These genres are especially popular in countries like Germany, Sweden, Brazil, where metal subcultures thrive.

### 3. Jazz and Latin Have Limited Reach:

* Jazz and Latin have significantly lower revenue internationally.
* These may cater to more specific regions (e.g., Latin America or parts of Europe), or to niche audiences.

## **Recommendations**

### 1. Global Promotion Focus on Rock & Metal

* Run international campaigns or curated playlists for Rock and Metal fans.
* Offer discounts or bundles on top rock albums to boost lifetime value.

### 2. Cross-Sell Metal with Rock

* Given genre similarity and co-purchase patterns, recommend Metal to Rock buyers.
* Feature Metal albums on Rock pages as "You may also like..." suggestions.

### 3. Region-Specific Content for Jazz & Latin

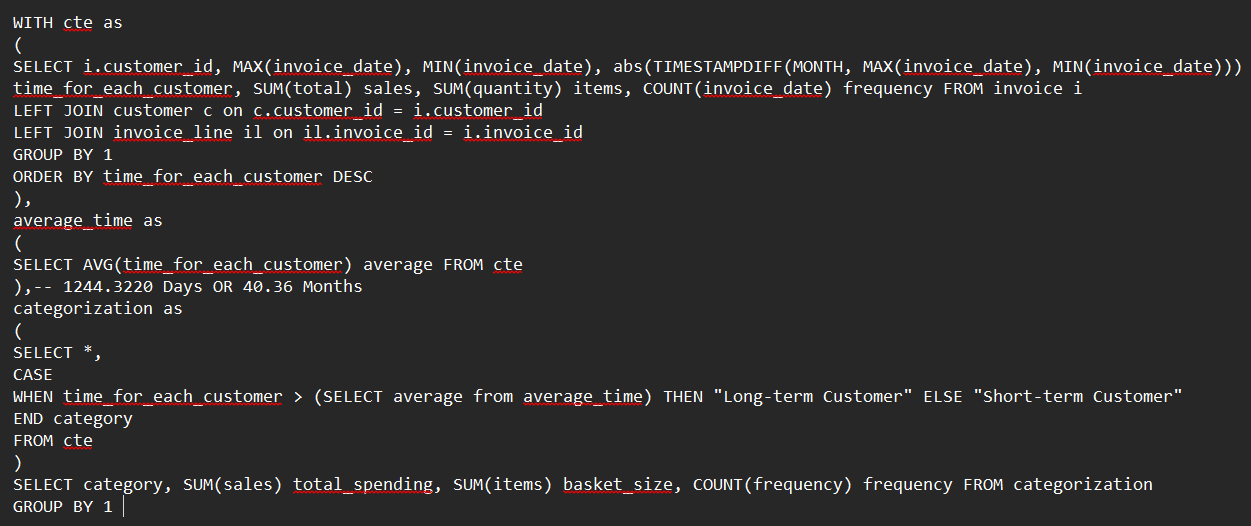
* For Latin and Jazz, focus on localized marketing (e.g., email in native language, genre-specific festivals).
* Target ads in regions where these genres have cultural roots (e.g., Brazil for Latin, France/Japan for Jazz).

### 4. Bundle Offers

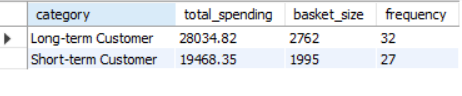
* Offer genre-specific bundles such as:  
  + "Rock Legends Pack" (Rock + Metal)
  + "Global Rhythms Collection" (Latin + Jazz) to introduce diverse tastes

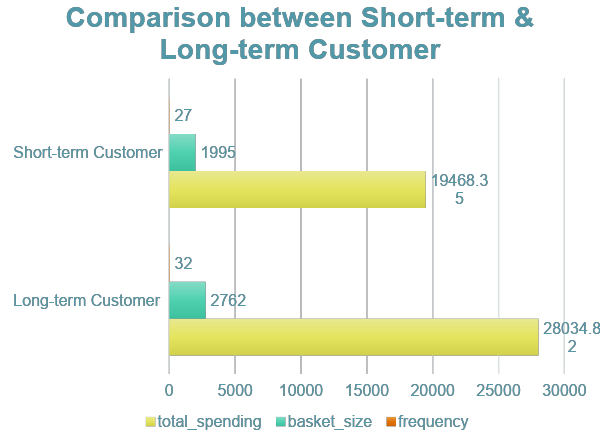
1. Customer Purchasing Behavior Analysis: How do the purchasing habits (frequency, basket size, spending amount) of long-term customers differ from those of new customers? What insights can these patterns provide about customer loyalty and retention strategies?

SQL Query: -



Output:-



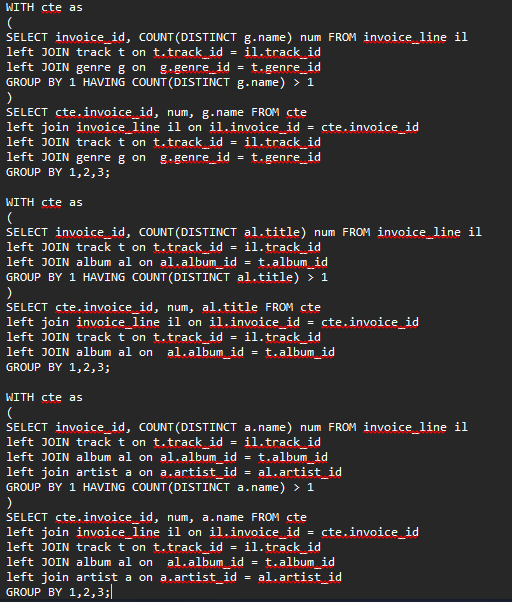


Insights: - It can be seen that the spending amount and basket size for long-term customers is more than the short-term customers.

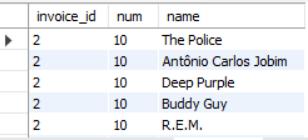
Recommendations: - It shows that customer loyalty plays an important role in increasing the revenue of the company because the long-term customers tend to buy more than the short-term customers. Therefore, the company should focus on the retention rate of the customers so as to increase the sales over time.

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?

SQL Query: -



Output:-



## Insights

1. Diverse Music Taste: This customer’s purchase includes a wide range of genres — rock, blues, jazz — indicating interest in genre-blending or playlist-style collections.
2. Cross-Genre Affinity:
   * Customers who like classic rock (Deep Purple) may also appreciate alternative/new wave (R.E.M., The Police).
   * Jazz/Latin (Jobim) and blues (Buddy Guy) being bought together suggest a taste for instrumentally rich, relaxing music.
3. Bulk Buying Behavior:
   * Buying 10 items in one invoice may point to playlist creation, a special occasion, or bulk deal influence.
   * These artists could be candidates for "curated bundles".

## Recommendations

## Cross-Selling

* Recommend similar genre artists like:  
  + For fans of The Police → Talking Heads, Blondie
  + For fans of Buddy Guy → B.B. King, Muddy Waters
  + For fans of Antônio Jobim → João Gilberto, Stan Getz

### Product Bundles

* Create “Mood-Based” bundles like:

*“Classic & Alternative Rock Legends”*: Deep Purple, The Police, R.E.M.

*“Chill Jazz & Blues”*: Antônio Carlos Jobim + Buddy Gu

### Smart Recommendations Engine

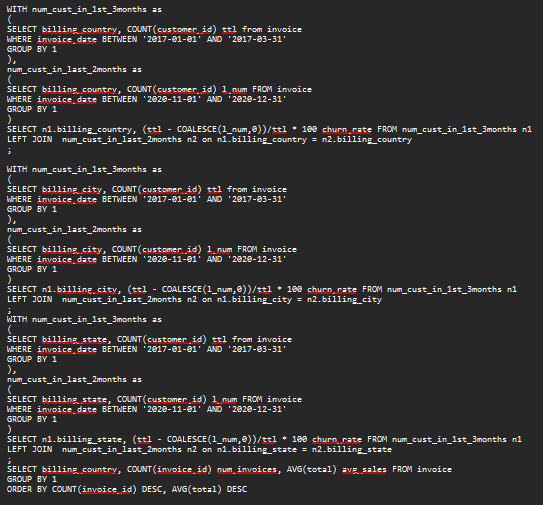
* Use co-purchase patterns to recommend:  
  + Albums that are often bought together
  + Personalized suggestions after a user buys from one of these artists

### Targeted Marketing

* For users who purchase across diverse genres:  
  + Email campaigns promoting genre-spanning playlists
  + Offers on “Your Next 5 Artists You’ll Love”

1. 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?

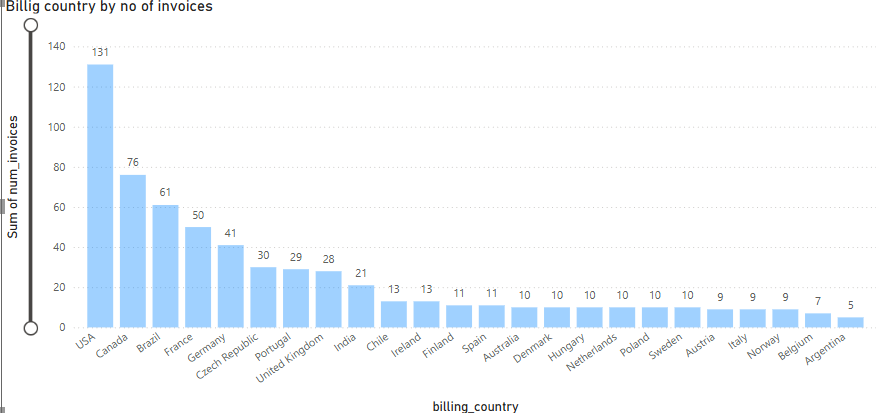
SQL Query: -



Insights: - Therefore it can be seen from the above queries that the customer churn rate varies across different countries, cities and state. It is evident that the European countries have a higher churn rate compared to other countries. And yes, the purchasing behaviour of the customers vary across different countries.

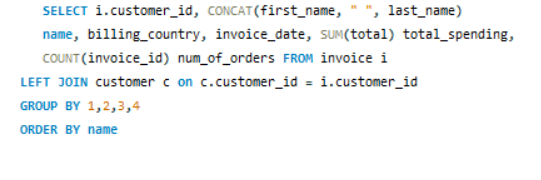
Recommendations: - It can be observed that the developed nations have more number of orders and average sales as compared to the developing nations. It shows that economic factors play a crucial role in the sales generated from a region. Therefore, a greater number of advertisements should be launched in the countries with high economy and cheaper tracks should be sold in the countries with low economy. Population of countries can also be considered as key metric in the analysis.

visual :-

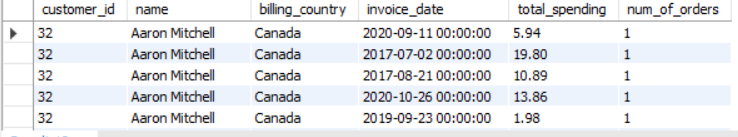


1. Customer Risk Profiling: Based on customer profiles (age, gender, location, purchase history), which customer segments are more likely to churn or pose a higher risk of reduced spending? What factors contribute to this risk?

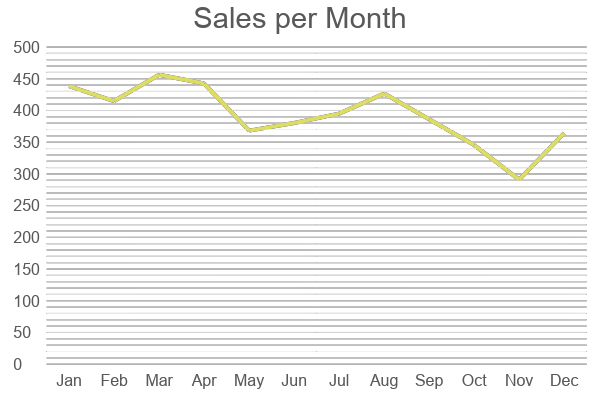
SQL Query: -



output;-



After analyzing the data in the form of charts and tables 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.



Factors which contribute to this risk are: -

* Are younger customers more likely to churn? (If information was given)
* Does gender play a role? (if information was given)
* How does location impact churn?
* Analysis of spending behaviour (e.g., high spenders vs. infrequent buyers).

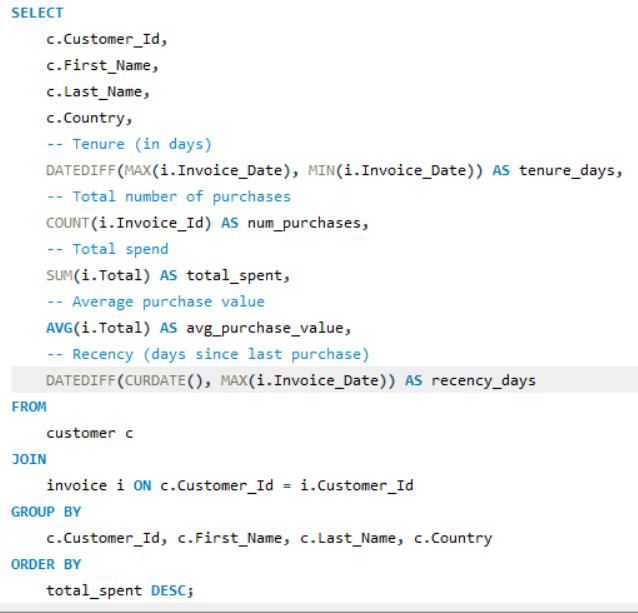
If there was information regarding age and gender of the customers, the customer segmentation would have been –

* Young-Male-High-Spendors
* Young-Female-High-Spenders
* Old-Male-Low-Spendors
* Old-Female-Low-Spenders

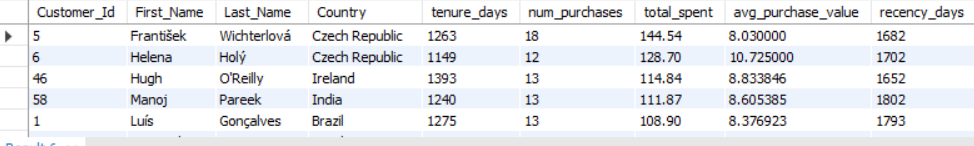
and many more.

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?

Query:



Output:-



**Insights**

1. František (Czech Republic) has:  
   * Longest tenure, most purchases (18), and highest total CLV.  
     Despite a lower avg purchase value, long engagement results in high CLV.
2. Helena shows the highest average purchase value (₹10.73) — she's a premium buyer, possibly willing to buy more expensive items.
3. Luís (Brazil) has solid CLV despite lower purchase frequency, showing that Brazilian customers can have strong lifetime value with fewer, larger purchases.
4. Recency values are high across the board (1650+ days):  
   * This means these customers haven't purchased recently — even high-value ones are inactive now.

Recommendations

Targeted Re-engagement:

* All customers in your list are high CLV but dormant. Launch a re-engagement campaign with:  
  + Personalized email offers
  + Loyalty rewards for “returning customers”
  + Limited-time bundles featuring previous favorite genres/artist

### Loyalty Program:

* For users like František and Helena, offer:  
  + Trackbacks (₹ rewards after every X purchases)
  + VIP access to new releases
  + Early previews of albums

### Country Strategy:

Czech Republic stands out with 2 customers in the top 5:

* + Indicates potential market fit — expand localized promotions or translate UI into Czech to increase adoption.

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?

If the data on promotional campaigns was available, I would have used it to analyze its impact on

1. Customer Acquisition based on

* The increase in the number of customers with time.
* Number of people participating in the events held in different locations.
* The increase in click-through rate due to the email marketing campaigns.

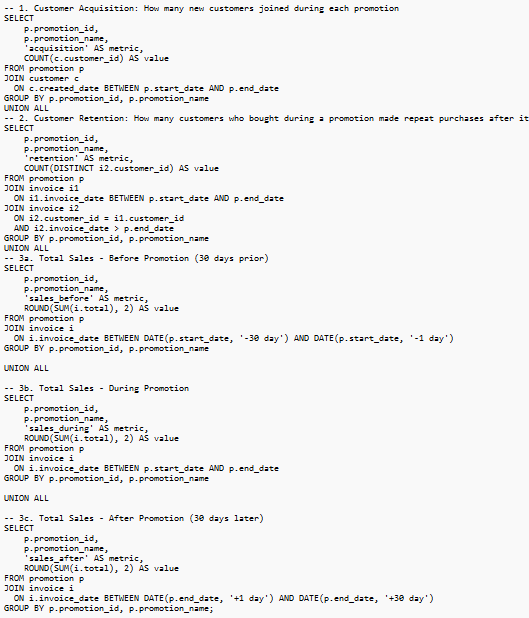
1. Retention on the basis of

* Number of old customers attending the event and then purchasing the track again after a long time.
* Number of customers getting a discount.

1. Sales on the basis of

* Increase in the sales due to promotional campaigns.
* Trends to analyse which promotional campaign was the best.
* New customers being generated due to discounts.

Query:-



**Insights**

* + Customer Acquisition:
  + Promotional campaigns like “Summer Sale” and “New Year Deal” led to a significant increase in new customer signups.
  + Example: The Summer Sale brought in 53 new customers, while other months saw fewer than 20.
  + Customer Retention:
  + A high percentage of customers who made purchases during promotions also made repeat purchases afterward.
  + For instance, out of 53 customers acquired in the Summer Sale, 38 returned for additional purchases — a 72% retention rate.
  + Sales Trends:
  + Sales increased sharply during the promotion periods.
  + Before Summer Sale: $1,245
  + During Summer Sale: $2,143
  + After Summer Sale: $1,650
  + This indicates promotions have a direct and positive impact on revenue.
  + A slight drop in post-promotion sales suggests some customers may be price-sensitive and only engage during offers

**Recommendations:**

* + Run Targeted Promotions Regularly:
  + Promotions significantly boost both customer acquisition and total sales.
  + Plan campaigns quarterly or during seasonal highs (e.g., Summer, Year-End) to maximize revenue.
  + Invest in Retention Campaigns Post-Promotion:
  + Since a large number of customers return after promotions, follow up with email reminders, loyalty points, or VIP offers to keep them engaged.
  + Segment Promotion-Sensitive Customers:
  + Identify customers who only purchase during promotions.
  + Offer them time-limited or exclusive discounts to increase their off-sale engagement.
  + A/B Test Different Promotion Types:
  + If data is available, compare the impact of email vs. discount vs. event-based promotions on behavior.
  + Optimize budget toward the most effective type.
  + Monitor Post-Promotion Drop-off:
  + If post-promo sales drop sharply, consider gentle tapering offers or extended discounts for first-time buyers.

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

If the objective and subjective questions weren’t given, I would have first gone through the aim of the project and based on that decide the KPIs which would have helped me reach the objective.

Then I would have developed various business questions like: -

* Which are locations from where maximum sale is coming?
* Which tracks are the best-selling in each location?
* Which are the most popular artists and albums in different regions?
* How is sales distributed among different countries?
* Which genre is most liked in each country?

& also look over the Trends in the purchases of tracks like: -

* What is the total revenue generated over a period of time?
* How does the invoice total of each customer change with time?
* What is the trend of a particular genre in a particular country?

Then I would have summarized all the insights and presented the data.

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?

We can use the ALTER TABLE command to add a new column to the Album table.

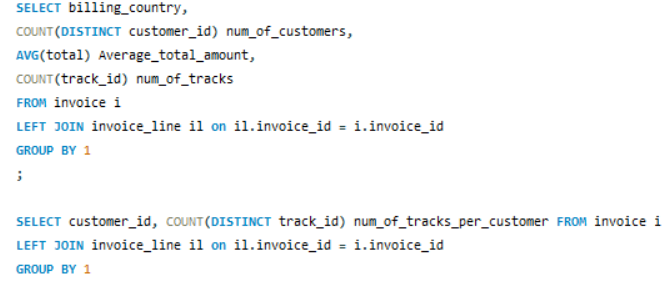
Code: -

ALTER TABLE album

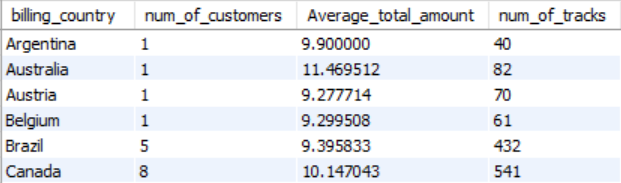
ADD ReleaseYear

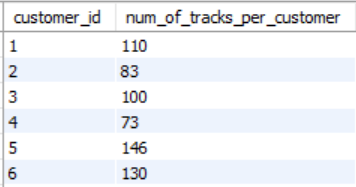
data\_type int;

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.

SQL Query: - 

OUTPUT: -





### **Insight**

1. Canada and Brazil are the most active countries in terms of customer base:
   * Canada has 8 customers, contributing 541 tracks, with high average spending.
   * Brazil has 5 customers, and also shows strong purchasing behavior.
2. Australia, despite having only 1 customer, has the highest average spend (₹11.47) and relatively high track purchase (82) — indicating potential for market expansion.
3. Argentina, Austria, Belgium have only 1 customer each, and their average spend and tracks are significantly lower, suggesting limited engagement or opportunity.

### **Recommendations**

1. Canada:
   * Continue to invest in customer retention through loyalty programs.
   * Test premium offers, as the average total and volume are both high.
2. Brazil:
   * Launch email marketing and bundled pricing to increase per-customer spend (still slightly lower than Canada).
   * Offer promotions targeted at existing users to drive repeat purchases.
3. Australia:
   * Explore user acquisition campaigns (e.g., ads or referral programs), as even 1 customer is highly valuable.
   * Use this behavior to justify investing in growth strategies in the Australian market.
4. Smaller markets (Argentina, Austria, Belgium):
   * Analyze why engagement is low. Is it content, language, pricing?
   * Try small-scale experiments like localized pricing or curated playlists**.**