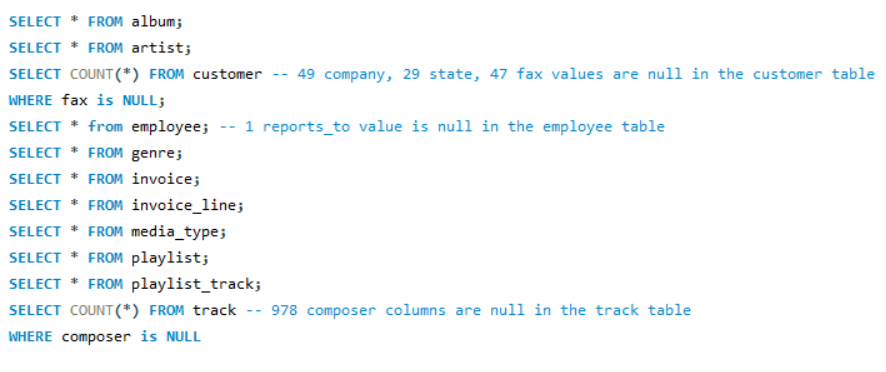
**OBJECTIVE QUESTIONS**

**Q1.** Does any table have missing values or duplicates? If yes, how would you handle it?

**ANS.** SQL QUERY:-

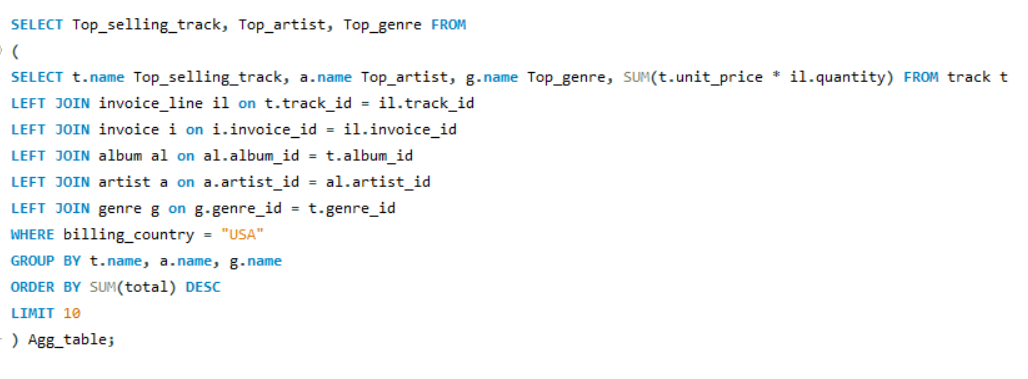
Yes, there were several missing values in the dataset, such as:

* **Customer Table**: 49 null values in the "Company" column, 29 in the "State" column, and 47 in the "Fax" column.
* **Employee Table**: 1 null value in the "Reports To" column.
* **Track Table**: 978 null values in the "Composer" column.

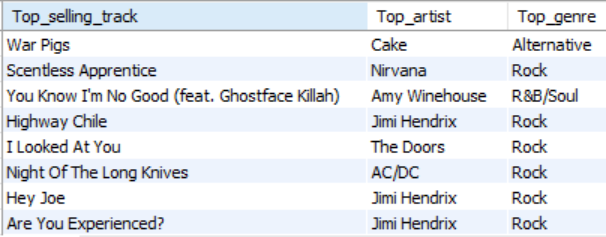
The dataset contains no duplicate values, ensuring data accuracy and reliability. The COALESCE function will be used to handle null values.

**Q2.** Find the top-selling tracks and top artist in the USA and identify their most famous genres.

**ANS.** SQL QUERY:-



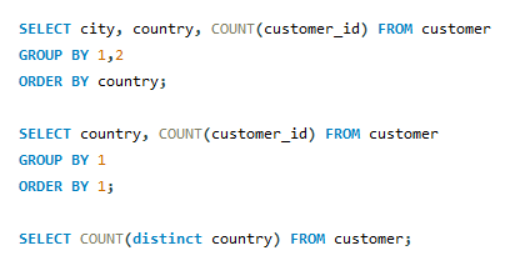
OUTPUT: -



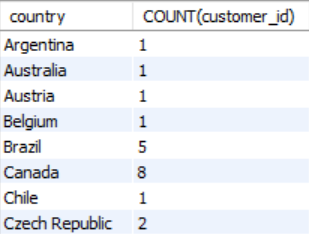
Therefore, the top genre in USA is Rock.

**Q3.** What is the customer demographic breakdown (age, gender, location) of Chinook's customer base?

**ANS.** SQL QUERY:-

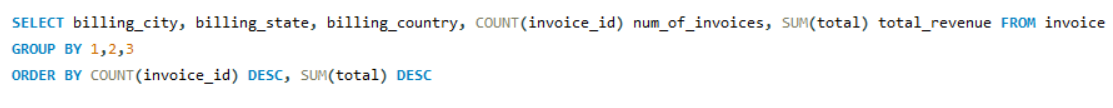


OUTPUT: -

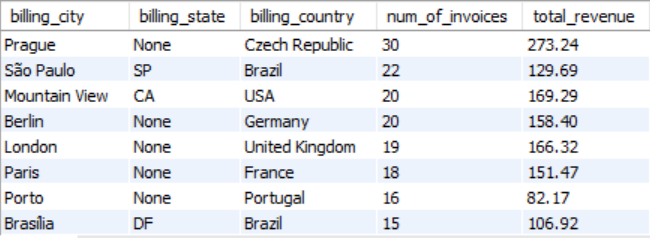
 

* The customer demographic breakdown based on location is highly diversified.
* Chinook's customer base spans 24 countries, with the highest number of customers from the USA.
* The customer table lacks age and gender columns, which makes it difficult to analyze the customer demographic breakdown in detail.

**Q4.** Calculate the total revenue and number of invoices for each country, state, and city:

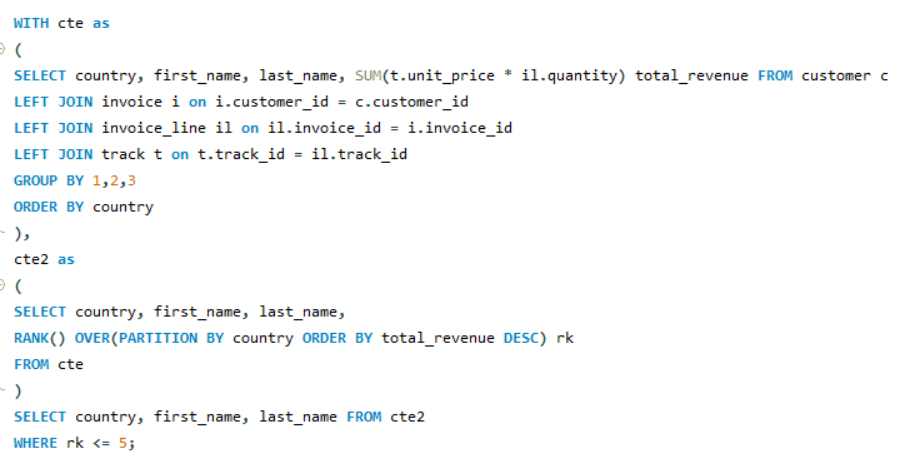
**ANS.** SQL QUERY:-

OUTPUT: -

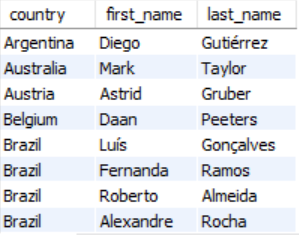


**Q5.** Find the top 5 customers by total revenue in each country.

**ANS.** SQL QUERY:-

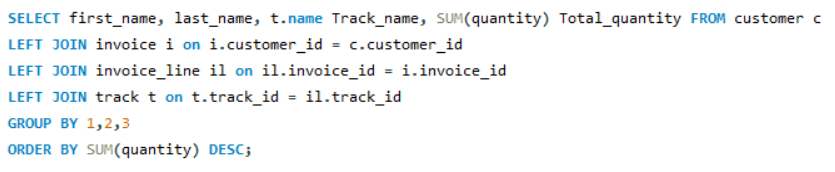


OUTPUT: -



**Q6.** Identify the top-selling track for each customer.

**ANS.** SQL QUERY:-



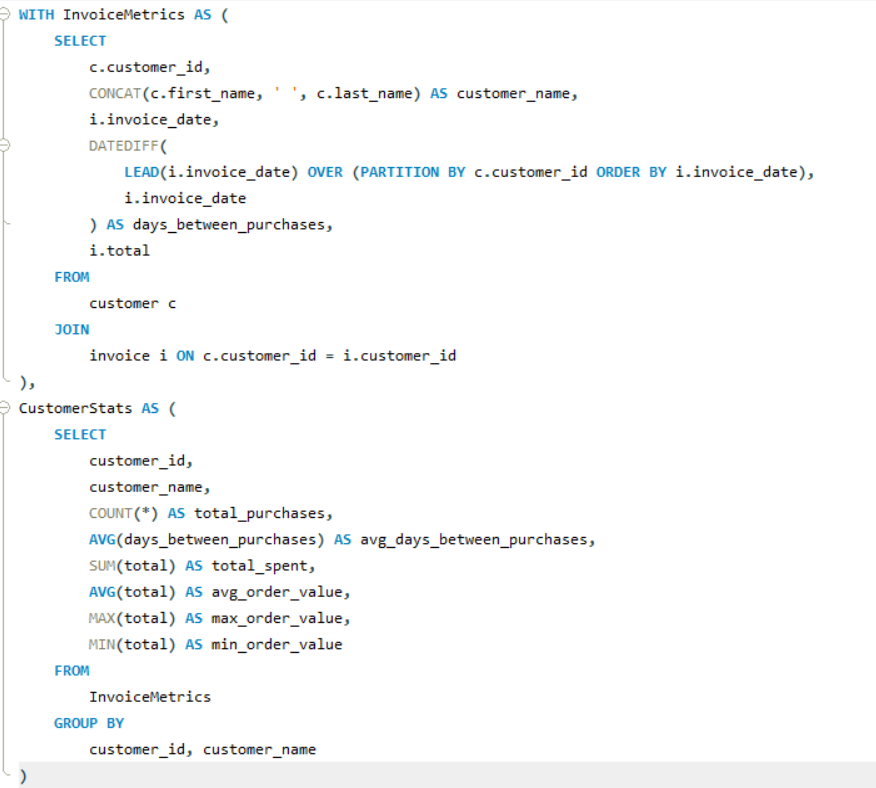
OUTPUT: -

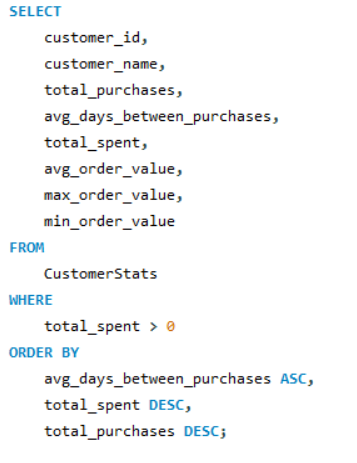


\*The customers purchased a similar quantity of different tracks, making it challenging to identify a single top track for each customer.

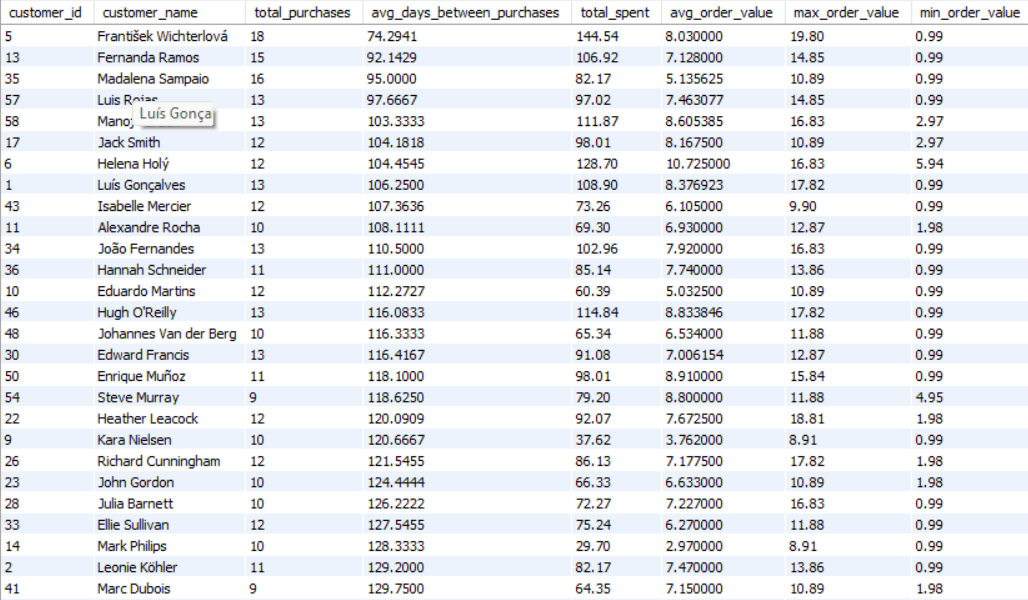
**Q7.** Are there any patterns or trends in customer purchasing behaviour (e.g., frequency of purchases, preferred payment methods, average order value)?

**ANS.** SQL QUERY:-



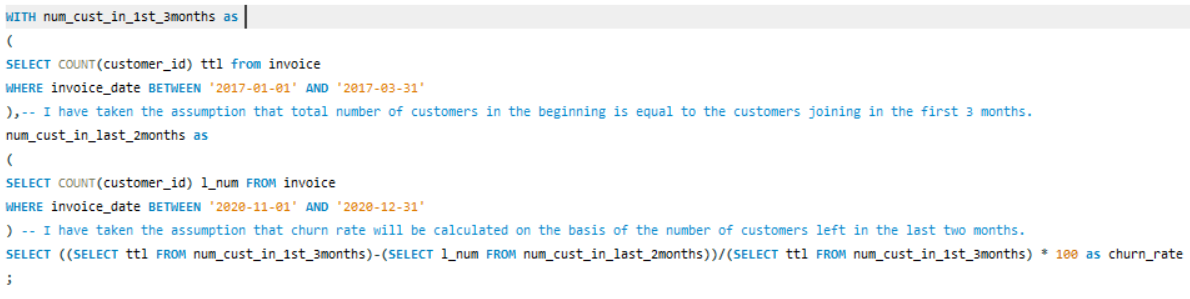


OUTPUT: -



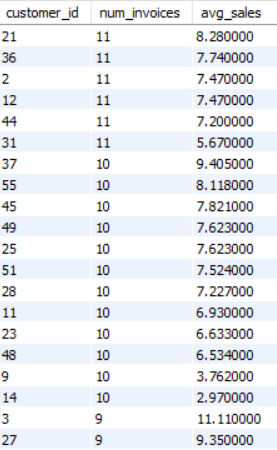
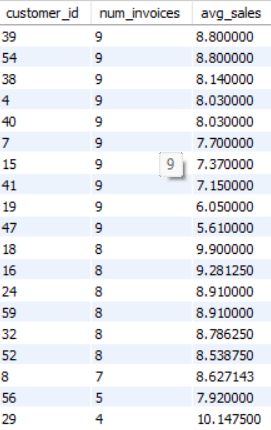
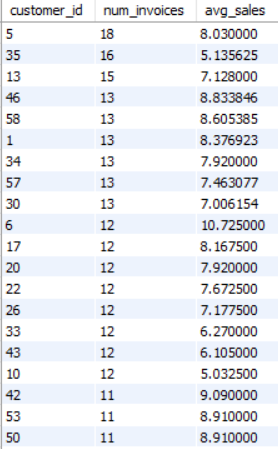
The table provides insights into customer behaviour, showing purchase frequency, total spending, and average order values. "Fernanda Ramos" and "František Wichterlová" are frequent buyers, with "František Wichterlová" also spending the most. Customers like "Luis Gonçalves" spend more per transaction, indicated by higher average order values. The avg\_days\_between\_purchases shows some customers have longer gaps between purchases, suggesting lower engagement. Customers like "Mano" have shorter gaps and higher purchase frequency. Focus on high-spending, frequent buyers, and use upselling or rewards to enhance loyalty.

**Q8.** What is the customer churn rate?

**ANS.** SQL QUERY:-

Customer churn rate is the percentage of customers who discontinue their business relationship with an organization over a specified period. It is a critical metric for evaluating customer retention and the overall health of a business.

OUTPUT: -

.

The customer churn rate of the company is **40.8163%**, calculated based on the total number of customers in the the first three months (49) and the number of customers remaining in the last two months (29).

The number of customers lost is:  
**49 - 29 = 20**

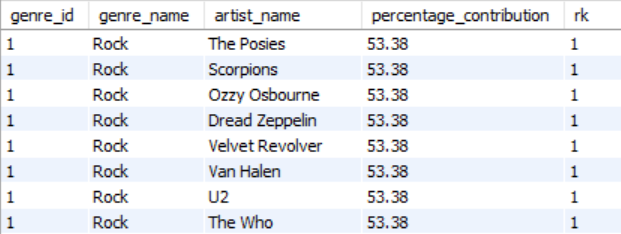
Churn rate = 20/49 = 40.8163

**Q9.** Calculate the percentage of total sales contributed by each genre in the USA and identify the best-selling genres and artists.

**ANS.** SQL QUERY:-



OUTPUT: -



Therefore the top selling genre in USA is Rock.

And

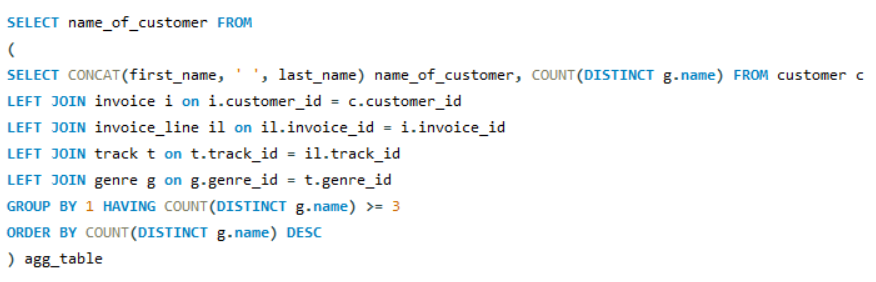
The following are the top artists associated with the Rock genre:

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, and Accept.

These artists represent a diverse range of styles and influences within the Rock genre.

**Q10.** Find customers who have purchased tracks from at least 3 different genres.

**ANS.** SQL QUERY:-

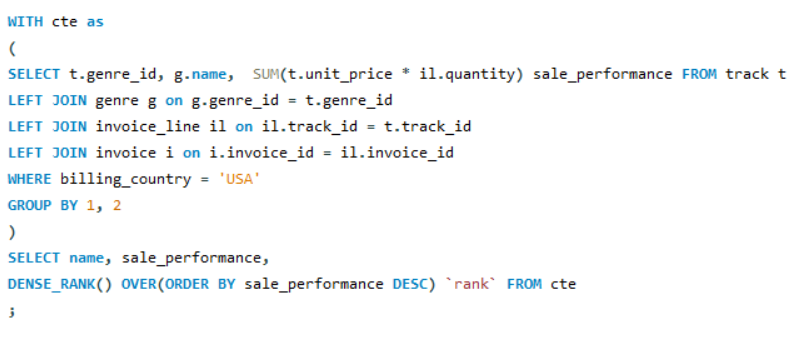


OUTPUT: -

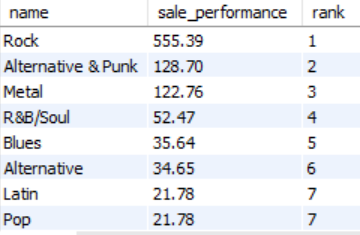
Leonie Köhler is a unique customer who has purchased tracks spanning 14 different genres, showcasing a diverse taste in music.

**Q11.** Rank genres based on their sales performance in the USA.

**ANS.** SQL QUERY:-



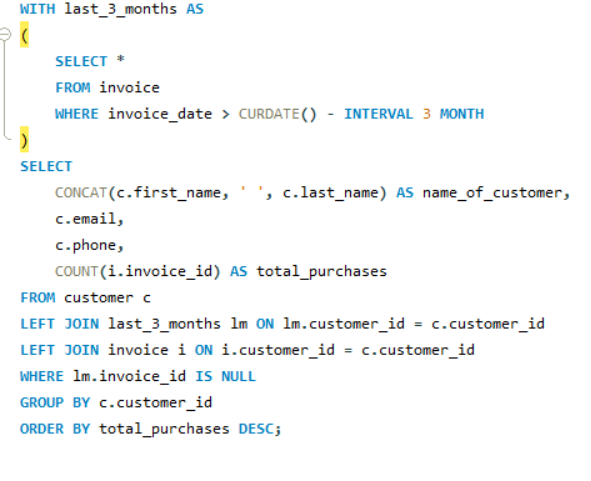
OUTPUT: -



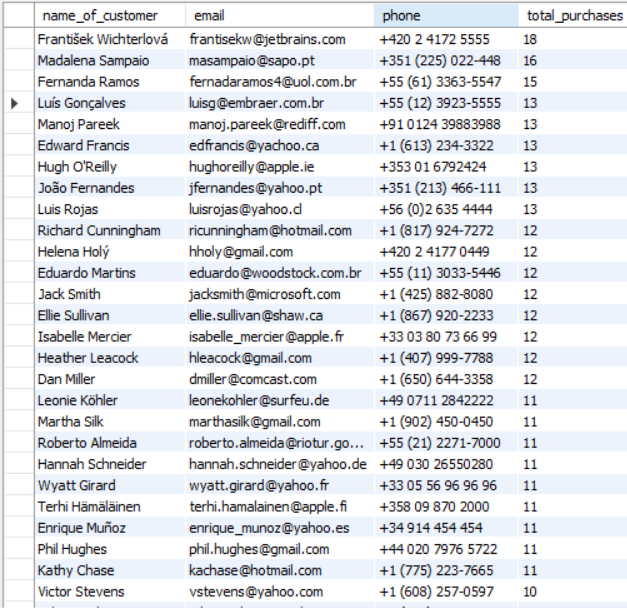
The sales performance of various music genres in the USA reveals a clear dominance of **Rock**, with a substantial lead at 555.39 USD, followed by **Alternative & Punk** and **Metal**. Genres like **R&B/Soul**, **Blues**, and **Alternative** also perform moderately well, but their sales are significantly lower than Rock. **Latin** and **Pop** share the same sales figure, ranking 7th, while genres like **Hip Hop/Rap** and **Jazz** show more modest performance. At the bottom, **Reggae**, **Electronica/Dance**, **Classical**, **Heavy Metal**, **TV Shows**, and **Soundtrack** have very low sales, indicating limited market traction. Overall, the rankings demonstrate a clear divide in genre popularity, with Rock far outperforming others.

**Q12.** Identify customers who have not made a purchase in the last 3 months.

**ANS.** SQL QUERY:-



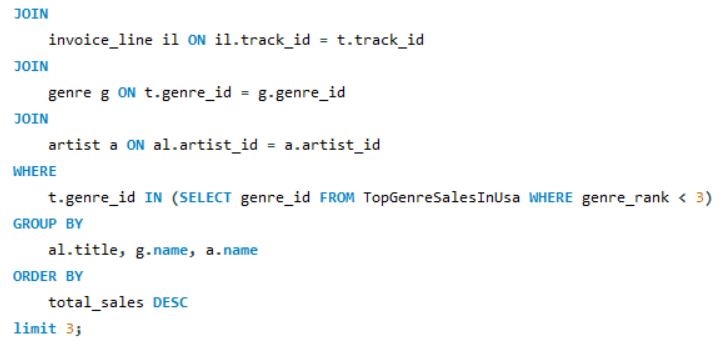
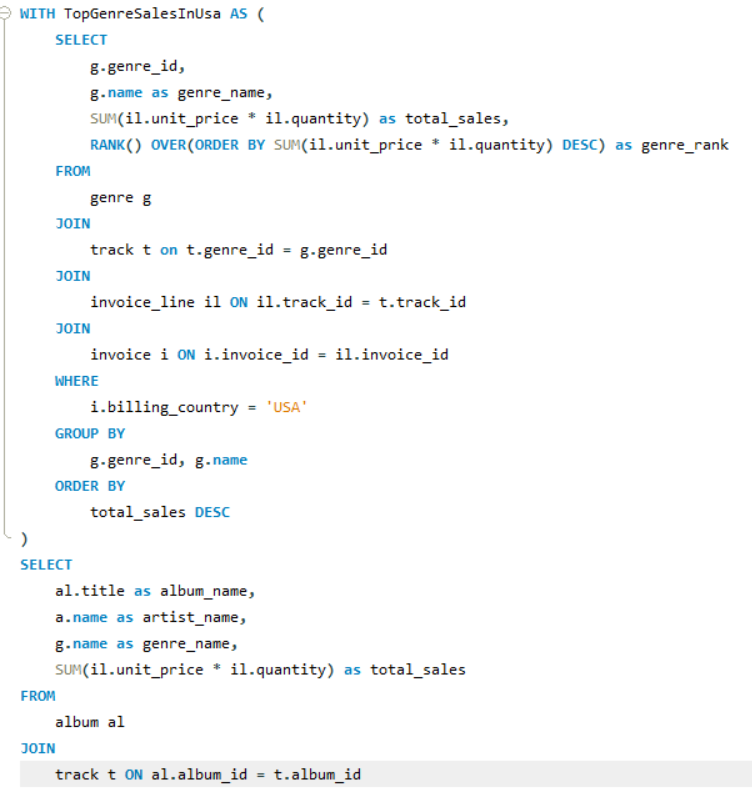
OUTPUT: -



**SUBJECTIVE QUESTIONS**

**Q1.** 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.

**ANS.** SQL QUERY:-



OUTPUT:-



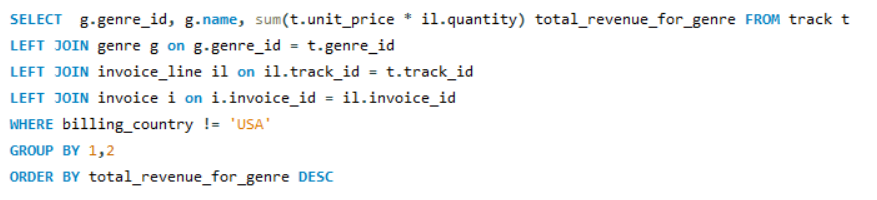
Recommendation: The top 3 albums that should be prioritised for advertisements and promotions in the USA based on genre analysis are:-

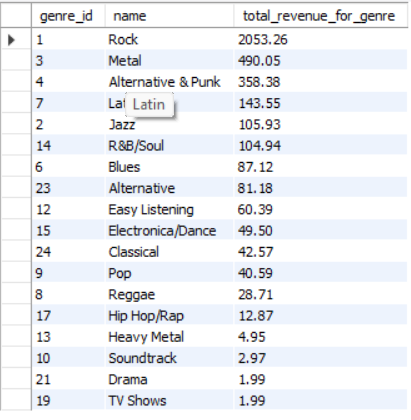
1. **Are You Experienced?** (By Jimi Hendrix)
2. **Get Born** (By JET)
3. **The Doors**  (By The Doors)

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

**Q2.** Determine the top-selling genres in countries other than the USA and identify any commonalities or differences.

**ANS.** SQL QUERY:-



OUTPUT: -

The commonality between the data for the USA and other countries is that the ROCK genre consistently holds the top spot in both. The 2nd and 3rd spots are occupied by the METAL and ALTERNATIVE & PUNK genres, respectively, across the world.

**Q3.** 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.** SQL QUERY:-



A screenshot of a computer screen

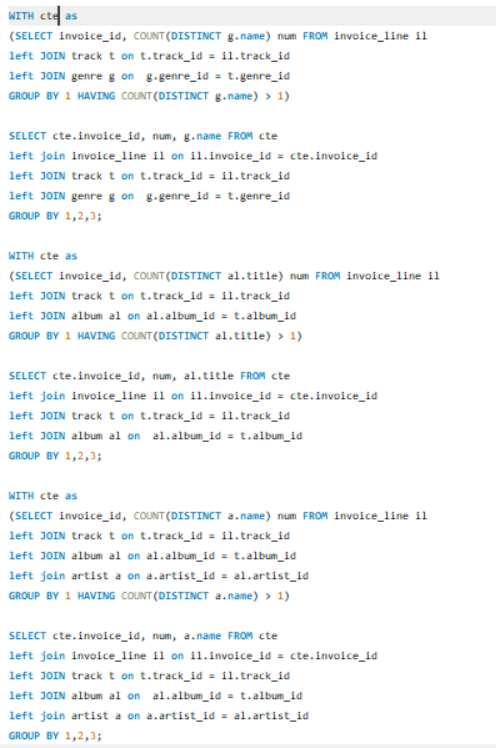
Description automatically generatedOUTPUT:-

**Insights**: - It can be observed that long-term customers tend to have higher spending amounts and larger basket sizes compared to short-term customers.

**Recommendations:** - This indicates that customer loyalty is crucial for boosting company revenue, as long-term customers tend to make more purchases than short-term customers. Therefore, the company should prioritize improving customer retention to drive higher sales over time.

**Q4.** 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.** SQL QUERY:-



OUTPUT:-

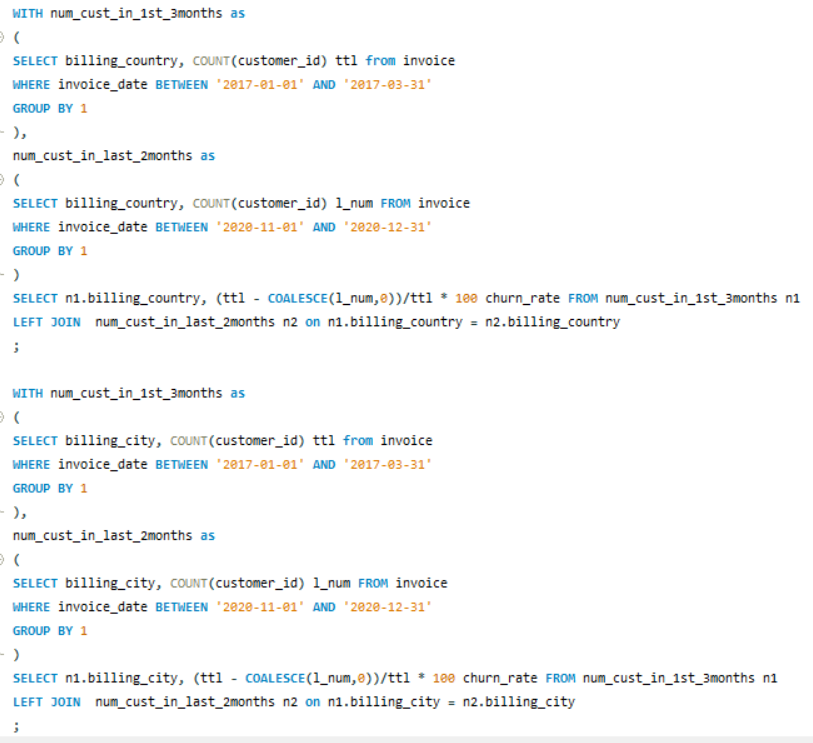
When the output of the query is analyzed in Excel through a Pivot Table, the following patterns emerge:

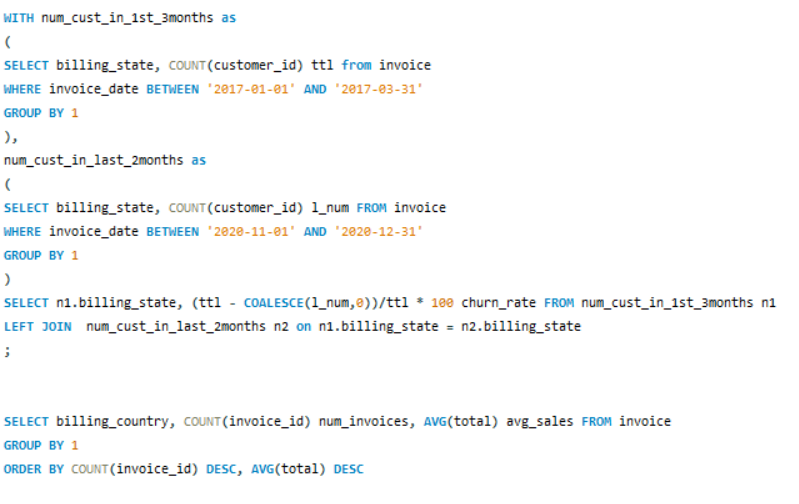
* **Genres**: Rock, Metal, and Alternative & Punk are the genres most frequently purchased together.
* **Albums**: Mezmerize, The Doors, and Dark Side of the Moon are the albums often purchased together.
* **Artists**: Green Day, Foo Fighters, and U2 are the artists frequently purchased together.

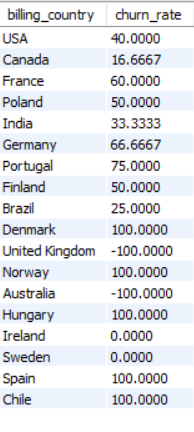
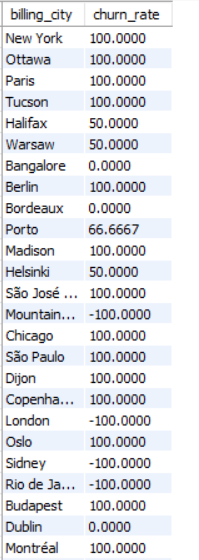
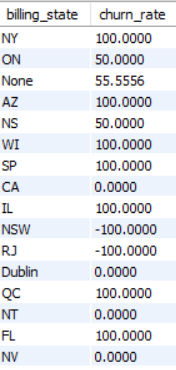
These insights highlight customer preferences and can guide bundling or recommendation strategies.

**Q5.** 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.** SQL QUERY:-





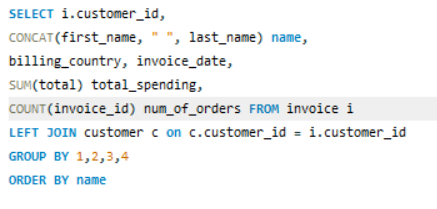
OUTPUT:-

Insights: - From the above queries, it is evident that the customer churn rate varies across different countries, cities, and states. European countries exhibit a higher churn rate compared to other regions. Additionally, customer purchasing behavior significantly differs across countries, reflecting diverse preferences and engagement levels.

Recommendations: - Developed nations have a higher number of orders and average sales compared to developing nations, indicating that economic factors significantly influence sales in a region. To optimize revenue, more advertisements should target high-income countries, while offering cheaper tracks in low-income nations. Additionally, population size can serve as a key metric for refining the analysis and tailoring strategies to specific markets.

**Q6.** 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.** SQL QUERY:-



OUTPUT:-

A screenshot of a computer screen

Description automatically generated

After analyzing the data through charts and tables, it is evident that countries with already high spending and order frequency are showing growth, while sales and frequency remain stagnant in other regions. This highlights the need for new promotional campaigns in underperforming countries to reduce the churn rate and sustain or increase spending levels.

Factors contributing to customer churn risk include:

* **Age**: Are younger customers more likely to churn? (If age data were available, it could provide insights.)
* **Gender**: Does gender influence churn rates? (Gender information could reveal trends if provided.)
* **Location**: How does geographical location impact churn? Regional differences may indicate varying engagement levels.
* **Spending Behavior**: High spenders versus infrequent buyers—does spending behavior correlate with churn risk?

Addressing these factors through targeted strategies could help mitigate churn effectively.

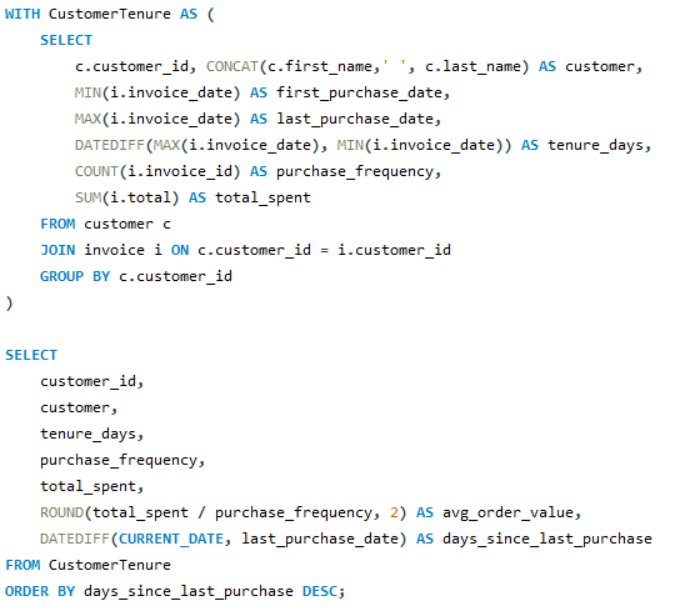
If age and gender data were available, customer segmentation could be more detailed, such as:

* **Young-Male-High-Spenders**
* **Young-Female-High-Spenders**
* **Old-Male-Low-Spenders**
* **Old-Female-Low-Spenders**
* And additional combinations reflecting varying spending patterns and demographics.

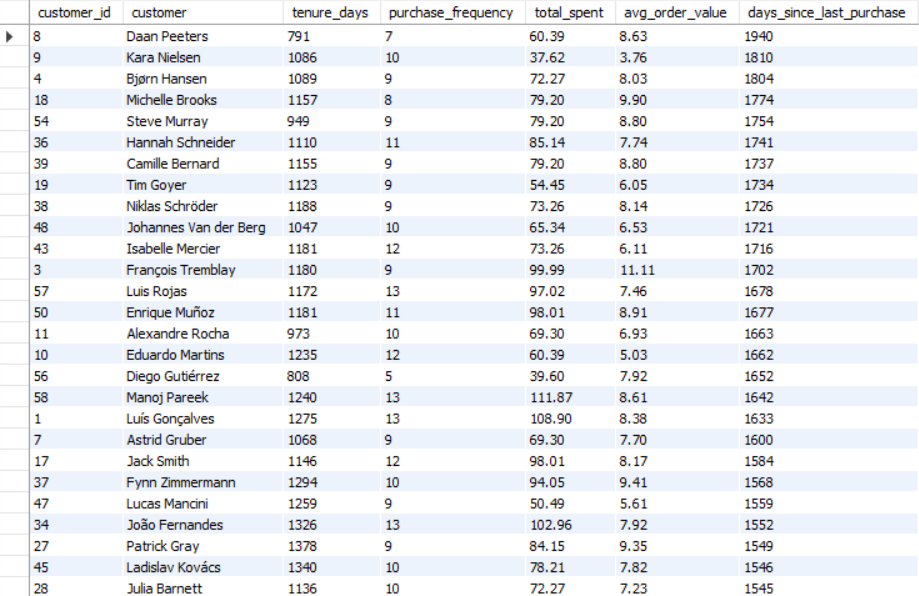
This segmentation would enable more targeted marketing and personalized strategies to optimize customer engagement and retention.

**Q7.** 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?

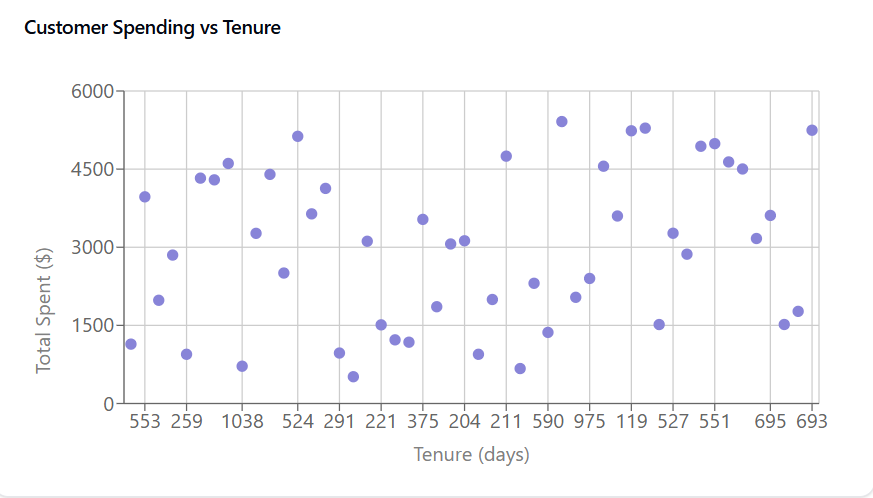
**ANS.** SQL Query:-

****

OUTPUT: -



1. To predict the lifetime value of a customer segment, analyzing purchase history and customer tenure is essential. These parameters help identify High-Value Customers who significantly contribute to revenue.
2. Customers with short tenure but high purchase activity should be targeted with promotional campaigns to convert them into loyal, long-term customers.
3. Customers who have stopped purchasing often belong to underdeveloped or developing countries, highlighting the influence of economic factors on sales. To address this, companies should leverage promotional channels like social media, articles, and advertisements, while offering discounts to boost sales in these regions.



**Q8.** 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.** If data on promotional campaigns were available, I would analyze their impact on:

1. **Customer Acquisition**:
   * Tracking the increase in customer count over time.
   * Monitoring participation in events across different locations.
   * Evaluating the rise in click-through rates from email marketing campaigns.
2. **Customer Retention**:
   * Identifying the number of returning customers who attended events and purchased tracks after a long gap.
   * Assessing the impact of discounts on retaining existing customers.
3. **Sales**:
   * Measuring the sales growth attributed to promotional campaigns.
   * Analyzing trends to identify the most effective campaigns.
   * Tracking new customer acquisition driven by discounts.

This would provide valuable insights into the effectiveness of promotional strategies.

**Q9.** How would you approach this problem, if the objective and subjective questions weren't given?

**ANS.** If the objective and subjective questions weren’t provided, I would begin by reviewing the project’s aim and identifying key performance indicators (KPIs) aligned with the objective. Based on this, I would develop relevant business questions, such as:

* Which locations contribute the most to sales?
* What are the best-selling tracks in each location?
* Who are the most popular artists and albums in different regions?
* How are sales distributed across countries?
* Which genre is most preferred in each country?

I would also analyze trends in track purchases, such as:

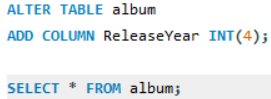
* Total revenue generated over time.
* Changes in the invoice total of each customer over time.
* Trends for specific genres in particular countries.

Finally, I would summarize these insights and present the findings comprehensively.

**Q10.** 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?

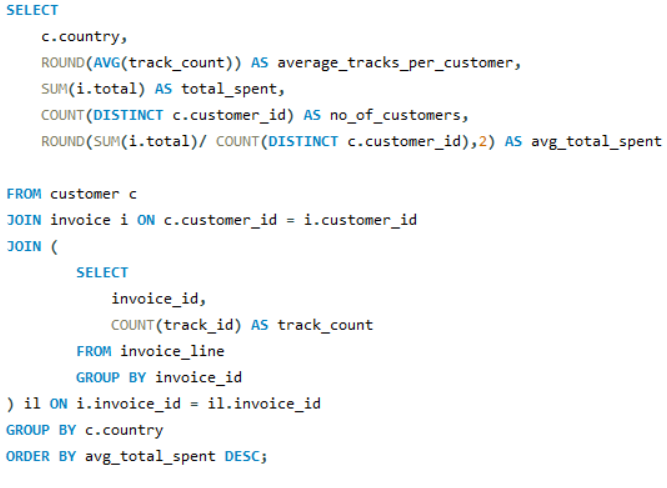
**ANS.** We can use the ALTER TABLE command to add a new column to the Album table.

SQL Query: -

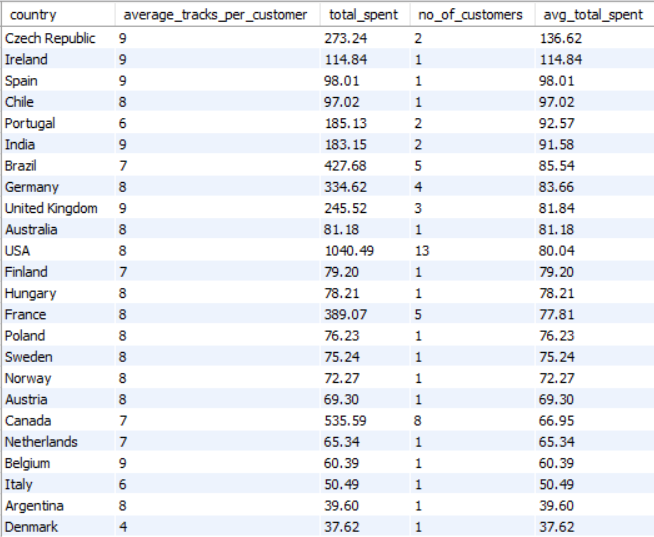


**Q11.** Chinook is interested in understanding the purchasing behavior of customers based on their geographical location. They want to know the average total amount spent by customers from each country, along with the number of customers and the average number of tracks purchased per customer. Write an SQL query to provide this information.

**Ans.** SQL Query: -



OUTPUT: -



The table reveals customer data across countries, showing metrics like the average number of tracks purchased, total spending, and the number of customers. The Czech Republic, Ireland, and Spain have high track averages, while the USA has the largest customer count (13) but lower average spending. Total spending varies, with Brazil and Canada showing higher values, while Chile and Finland have smaller totals. The USA leads in total spending at $10,040.49. Countries like Belgium and Italy have consistent average spending per customer at around $60.39. This data highlights regional differences in purchase frequency and spending behavior.